

Dissertation

**Developing a Consumer Health Informatics Decision**

**Support System Using Formal Concept**

**Analysis**

By

Vincent Horner

April 2007

Submitted in partial fulfillment of the requirements for a Master's degree (Computer Science) in the Faculty of Engineering, Built Environment and Information Technology, University of Pretoria, South Africa.

**Supervisor: Prof. D. G. Kourie**

**Co-supervisor: Dr. S. Obiedkov**

## Acknowledgements

I would like to thank the team of experts; Sr. Ndema, and Prof. vd. Merwe, and Prof. Mpe of Medunsa medical campus of the Limpopo University. Prof. Mpe was like a co-supervisor the generous way he gave to the project. The director of our unit, Dr. Nolwazi Mbanaga of the Medical Research Council of South Africa, whose original idea is the CDSS (Consumer Health Decision Support System). Lyn Hanmer my manager and mentor who battled together to bring about the CDSS and to conduct this study. To Sergei from Russia, who is my co-supervisor, who also became a friend for life while he was here in South Africa. To Prof. Kourie my supervisor for his guidance, patience and generosity. To Sr. Magolo and Sr. Mohlape who were my main contacts at Skinner Street Clinic where the patient data was collected. They helped beyond my initial requests, from your contact I have grown in the field of care of humans. To you all I am better now, technically and as a human being.

## Abstract

A consumer health decision support system (CDSS) is being developed at the South African Medical Research Council (MRC). It is a software program intended to help members of the public decide when they may be at risk of some common but serious illnesses like tuberculosis and hypertension. It would be ideal for a public health kiosk or e-health programs of the government. The program has been built as an expert system. Its knowledge base consists of rules which are used in assessing the risk of illness. The rules were given by medical experts who took part in the development of the CDSS. The study proposes a method for the evaluation of the rule base of the CDSS using FCA methods. It is important to evaluate the knowledge base of an expert system, because if its knowledge base is of broad scope and is accurate then it can be expected that the expert system will be good at giving advice and hence potentially useful. FCA is a mathematical framework which can be used to investigate causal relations in data. The study explored its utility in the evaluation of the CDSS knowledge base. FCA implications and the FCA formulation of the JSM method were two FCA methods that were selected. The FCA methods were used to generate rules from actual patient data, and these were compared to the rules initially given by the experts. The motivation to use FCA data analysis as well as experts' knowledge in the development of the CDSS program is that FCA data analysis may discover some things that the experts may have overlooked. Or at least the experts can review their expertise against actual field data which has been analysed by FCA methods. A system like the CDSS cannot be built using FCA data analysis techniques only, involvement of experts is very important. The two FCA methods were chosen so as to compare their results, and it was also thought that they may perhaps complement each other. Preliminarily it was found that FCA implications and the FCA formulation of the JSM method can be used in the evaluation of the rule base of the CDSS.

## TABLE OF CONTENTS

Declaration	0
Acknowledgements	i
Abstract	ii
List of Figures	ix
List of Tables	x
<b>CHAPTER 1: Introduction</b>	<b>1</b>
1.1 Project Description	1
1.2 Areas of Study	6
1.3 Study Proposition	6
1.4 Study Design	7
1.5 Data Collection	8
1.6 Summary of Results	9
1.7 Dissertation Outline	10
<b>CHAPTER 2: Consumer Health Informatics</b>	<b>14</b>
2.1 Introduction to Consumer Health Informatics	14
2.2 Origin of the CDSS Project	16
2.2.1 How the Project Came About	16
2.2.2 Health Problem Addressed	17
2.2.3 Criteria for Selecting Study Illnesses	18
2.3 The Medical Diagnosis Process	20

2.3.1	Background	20
2.3.2	Modeling Diagnostic Reasoning	21
2.3.3	CDSS Diagnosis Strategy	23
2.4	Ethical Issues of Decision Support	25
2.4.1	Standard View of Diagnosis	26
2.4.2	Care Standards	26
2.4.3	Appropriate Use and Users	26
2.4.4	Professional Relationship	26
2.4.5	Legislation of Decision Support Software	27
2.4.6	Problem of Illiteracy	27
2.4.7	Ethical Profile of CDSS	27
2.5	Conclusion	28
	<b>CHAPTER 3: Expert Systems (Decision Support Systems)</b>	<b>29</b>
3.1	Introduction to Expert Systems	29
3.1.1	Definition of an Expert System	29
3.1.2	History of Expert Systems	30
3.2	Knowledge Acquisition	31
3.3	Knowledge Representation	32
3.3.1	Production Rules	33
3.3.2	Networks	34
3.3.3	Objects	34
3.4	Explanation	35
3.5	Expert System Development Languages	35

3.5.1	CLIPS	35
3.5.2	LISP	36
3.5.3	PROLOG	36
3.5.4	Procedural Languages	36
3.6	Examples of Expert Systems	37
3.6.1	MYCIN a Classic Expert System	37
3.6.2	Recent Examples	38
3.7	Conclusion	38
<b>CHAPTER 4: The Design and Implementation of the CDSS</b>		<b>40</b>
4.1	Role of CDSS	40
4.2	Software Implementation of CDSS	41
4.2.1	CDSS Diagnosis Strategy	41
4.2.1	Software Modules of CDSS	43
4.3	User Interface Design	45
4.3.1	Criteria of CDSS User Interface	45
4.3.1	User Interface Design Challenges	46
4.4	Project Team	46
4.5	Project Evolution	47
4.6	Conclusion	48
<b>CHAPTER 5: Health Expert System Evaluation</b>		<b>49</b>
5.1	Introduction to Health Expert System Evaluation	49
5.1.1	Kinds of Evaluation	49
5.1.2	Factors in Health Expert System Evaluation	51

5.2	Evaluation Approach of MYCIN	52
5.2.1	Domain of MYCIN	52
5.2.1	MYCIN Evaluation Plan	53
5.2.3	MYCIN Evaluation Results	53
5.3	Evaluation Approach of Toxicology Analysis by Means of the JSM Method	54
5.3.1	Domain of Predictive Toxicology Challenge (PTC)	54
5.3.2	PTC Evaluation Plan	55
5.3.3	Results of PTC	56
5.4	Study Designs of Knowledge Base Evaluations	56
5.5	Evaluation Approach of CDSS Study	57
5.6	Conclusion	59
	<b>CHAPTER 6: Data Collection</b>	<b>60</b>
6.1	Source of Data	60
6.2	Patient Sampling	61
6.3	Formatting of Study Patient Data	64
6.4	Conclusion	65
	<b>CHAPTER 7: FCA Data Analysis Techniques</b>	<b>66</b>
7.1	Introduction to Formal Concept Analysis (FCA)	66
7.1.1	History of FCA	66
7.1.2	Definition of a Formal Context	67
7.1.3	Definition of the Single Prime Operation	68
7.1.4	Definition of a Concept	68

7.1.5	Definition of a Lattice	69
7.2	FCA Implications	71
7.2.1	Basic Definitions of FCA Implications	71
7.2.1	The Duquenne-Guigues Basis	74
7.2.3	Concept Explorer: Automated Tool for FCA Implications	75
7.3	JSM Hypotheses	76
7.3.1	Introduction to the JSM Method	76
7.3.2	Definition of a Hypothesis	76
7.3.3	Example	77
7.3.4	Choice of Intents as Hypotheses	79
7.3.5	QUDA: Automated Tool for JSM Hypotheses	80
7.4	Examples of Rules by JSM Hypotheses and FCA Implications	80
7.4.1	Rules by JSM Hypotheses	80
7.4.2	Rules by FCA Implications	80
7.5	Conclusion	81
	<b>CHAPTER 8: Data Analysis</b>	<b>82</b>
8.1	How to Interpret Results	83
8.2	Analysis of Tuberculosis	85
8.2.1	Experts' Rules for Tuberculosis	86
8.2.2	Training Dataset for Tuberculosis	86
8.2.3	FCA Implications for Tuberculosis	87
8.2.4	JSM Hypotheses for Tuberculosis	90
8.2.5	Summary of FCA Analysis for Tuberculosis	92

8.2.6	CDSS Performance for Tuberculosis	94
8.3	Analysis of Chronic Bronchitis	96
8.3.1	Experts' Rules for Bronchitis	96
8.3.2	Training Dataset for Bronchitis	96
8.3.3	FCA Implications for Bronchitis	97
8.3.4	JSM Hypotheses for Bronchitis	98
8.3.5	Summary of FCA Analysis for Bronchitis	99
8.4	Analysis of Hypertension	100
8.4.1	Training Dataset for Hypertension	100
8.4.2	FCA Implications for Hypertension	101
8.4.3	JSM Hypotheses for Hypertension	101
8.4.4	Summary of FCA Analysis for Hypertension	102
8.4.5	CDSS Performance for Hypertension	102
8.5	Summary of Data Analysis	104
8.6	Conclusion	105
8.6.1	Discussion of Results	105
8.6.2	Experts' Feedback Discussion	109
	<b>CHAPTER 9: Conclusion</b>	<b>111</b>
9.1	Summary of Study	111
9.2	Summary of Study Design	112
9.3	Review of Study Proposition	114
9.4	Summary of Data Analysis	114
9.5	Discussion: Consumer Health Informatics	116

9.6	Discussion: Expert Systems	118
9.7	Gaps	118
	<b>REFERENCE LIST</b>	<b>120</b>
	<b>APPENDIX I: Glossary of Terms</b>	<b>127</b>
	<b>APPENDIX II: CDSS Questionnaire</b>	<b>135</b>
	<b>APPENDIX III: Experts' Rules</b>	<b>139</b>

## LIST OF FIGURES

Figure 2.1	The diagnosis process.	17
Figure 2.2	CDSS and the diagnostic process.	17
Figure 2.3	Burden of disease index.	20
Figure 2.4	CDSS diagnosis strategy.	24
Figure 3.1	Knowledge model of an expert system.	32
Figure 4.1	CDSS and the diagnosis process.	40
Figure 4.2	CDSS diagnosis strategy.	42
Figure 4.3	CDSS and the model-control-view scheme.	43
Figure 4.4	CDSS diagnosis modules and their interaction.	44
Figure 4.5	Relationship of CLIPS with CDSS.	44
Figure 4.6	Screenshot of CDSS user interface.	45
Figure 7.1	Lattice of “Some Life and Water” context.	70

## LIST OF TABLES

Table 1.1	Number of study patients collected.	8
Table 1.2	Summary of comparison of FCA generated rules and experts' rules.	9
Table 1.3	Summary of CDSS performance.	9
Table 2.1	Illnesses selected for the CDSS.	18
Table 2.2	Summary of burden of disease index.	19
Table 5.1	Results of MYCIN evaluation.	54
Table 6.1	Number of study patients collected.	63
Table 6.2	Data collection schedule.	64
Table 6.3	Formatting of study patient data, example.	64
Table 7.1	Context: "Some Life and Water".	67
Table 7.2	Example context from Blinova et al (2003).	78
Table 8.1	How to interpret results, example context.	83
Table 8.2	Tuberculosis symptoms.	85

Table 8.3	Training dataset for tuberculosis.	86
Table 8.4	Experts' rules subsumed, tuberculosis.	93
Table 8.5	New rules, tuberculosis.	93
Table 8.6	Test dataset for tuberculosis.	94
Table 8.7	Performance of CDSS tuberculosis.	95
Table 8.8	Bronchitis symptoms.	96
Table 8.9	Training dataset for bronchitis.	97
Table 8.10	Experts' rules subsumed, bronchitis.	99
Table 8.11	New rules, bronchitis.	99
Table 8.12	Hypertension symptoms.	100
Table 8.13	Training dataset for hypertension.	100
Table 8.14	New rules, hypertension.	102
Table 8.15	Test dataset for hypertension.	102
Table 8.16	Performance of CDSS for hypertension.	103

Table 8.17	Summary of comparison of FCA implications and experts' rules.	104
Table 8.18	Summary of comparison of JSM hypotheses and experts' rules.	104
Table 8.19	Summary of comparison of FCA generated rules and experts' rules.	105
Table 8.20	Summary of CDSS performance.	105
Table 9.1	Summary of comparison of FCA generated rules and experts' rules.	115
Table 9.2	Summary of CDSS performance.	115

# Chapter 1

## Introduction

The outline of the first chapter is as follows. First a description of the project is made. This is a prelude which explains how the project came about, and describes the goals and the main issues addressed. Then the themes and areas of study are given. This is followed by a description of the study proposition and the study design. A summary of data collection and of results is made. Lastly is an outline of the dissertation. This chapter has not been extensively referenced, this is done in subsequent chapters where the themes introduced here are discussed in detail. The following are the headings of this chapter.

- 1.1 Project Description
- 1.2 Areas of Study
- 1.3 Study Proposition
- 1.4 Study Design
- 1.5 Summary of Data Collection
- 1.6 Summary of Study Results
- 1.7 Dissertation Outline

### 1.1 Project Description

This project is about the development and evaluation of a Consumer Health Informatics (CHI) software program called the Consumer Decision Support System (CDSS) for Important Common Ailments found in South Africa. Evaluation is an important part of development of a system, as it is used to identify weaknesses which can then be corrected. This results in an incremental process of development, whereby changes are made, and the revised system is evaluated, further weaknesses identified, improvements made, etc. Formal concept analysis (FCA) was used to evaluate the knowledge base of the CDSS. Before the discussion on the use of FCA an introduction of the CDSS is made, and then further on in this section FCA is introduced.

The CDSS program is intended to help members of the public (consumers / patients) decide when to seek professional medical help when there is a problem. This goal can be explained by describing its typical operating scenario. The archetypal user of the CDSS program is someone who may be experiencing some symptom (or symptoms) but who may not be sure what the

underlying problem may be. The person should not be gravely ill or urgently need to see a doctor. It is not expected that someone who is an urgent case should consult the CDSS program. Instead, such a person is expected to go to a primary health care (PHC) clinic or hospital immediately.

The scenario can be expanded as follows. Due to a lack of awareness of early signs of many common illnesses in South Africa patients tend to procrastinate in seeking help when they should. In many regions of South Africa, the scarcity of medical practitioners and the large distance to medical facilities exacerbates the problem. Therefore it was decided that ideally that illnesses to be selected for the study should be common and potentially harmful, but which can be readily treated if detected early. These illnesses should be part of those which have been identified to be problematic by the national department of health. The illnesses which were chosen are: tuberculosis, chronic bronchitis (smokers cough), hypertension, peptic ulcers and malaria. There are other illnesses which meet the study selection criteria but which have not been included in order to limit the scope of the study. When fully operational, the CDSS program may also include these other illnesses. Some examples of these are diabetes, asthma, and heart problems.

Access to health information is suggested by many research studies as a means to enable patients to be more active participants in the treatment process, leading to better medical outcomes (Eysenbach, 2000). This is the general principle that motivated for the CDSS program.

It operates by prompting the patient (consumer) to enter his or her symptoms. Thereafter it assesses the potential that the user may be suffering from one of the common ailments on its list, and then recommends an action. Its conclusion can be loosely considered to be a preliminary diagnosis. In the project it is loosely called a diagnosis. The use of diagnosis is not in the strict medical sense which has strong legal and ethical implications. The process to use the CDSS by the patient can be summarised as follows.

- The patient supplies symptoms to the CDSS program.
- The CDSS program suggests an illness from its common illnesses list which best explains the patient's symptoms.
- Finally it recommends an action to the patient. The recommendation of the CDSS is explained a little further below.

A field prototype of the CDSS program was developed and completed for use in the study. It was decided that the expert system style be used in its development. Expert systems have been used extensively in the past during the classic period of Artificial Intelligence in the nineteen seventies and nineteen eighties. They are still used today, though they do not have the same profile of those earlier years. It was found that many decision support systems have been built in the expert system style. The CDSS program has the characteristic of a decision support system, hence the choice of expert system for its development.

The domain experts on the development team included a chest specialist and a gastroenterologist, who are members of faculty at the Medunsa Medical campus of the University of Limpopo in South Africa. Also included as a domain expert, was a nursing sister with over thirty years of experience in primary health care. The team's knowledge engineer is the author of the dissertation.

The knowledge derived from experts was about the relationship between symptoms and illnesses. This knowledge was encoded as rules, which form part of the knowledge base of the CDSS program. An example of an expert rule is the following: {PC WL NS  $\rightarrow$  TB}. In ordinary language this states that persistent cough (PC), weight loss (WL) and night sweat (NS) strongly suggest tuberculosis (TB). The experts provided many rules of the type given above to cover all the illnesses chosen for the study.

However, the CDSS only includes rules which make *positive predictions*. Rules which make *negative predictions* are excluded. For example, the following rule is not allowed. "If the user has a persistent cough of only three days, and no other symptoms or risk factors then no TB is present." If similar negative conclusions are possible, then self treatment at home might be an option.

However, in the absence of negative predictions, the CDSS in its present form *always* recommends that the user should undergo further medical examination, either to confirm the suspected diagnosis at which the CDSS system has arrived, or to establish a diagnosis in the absence of a CDSS-derived one. Even in this preliminary form, it is considered it would still be useful. This is because it is assumed that most of its users do indeed have a problem – they have simply not yet decided to seek help. In practice, this tardiness in seeking help may be due to distance from medical facilities, lack of motivation, the all-to-human tendency to procrastinate, etc. Its role would be to give added encouragement to seek medical advice; to provide

information about a likely diagnosis where possible; perhaps to direct a user to a medical establishment that has specialized facilities for certain illnesses when necessary, etc.

Later it can perhaps be further developed to include positive and/or negative rules that predict minor illnesses (e.g. a hangover and sore throat), for which users may self-treat at home without the need to go to a PHC clinic. Clearly, however, such bypassing of medical opinion could be highly contentious and raise significant ethical questions. At this stage, therefore, there is no question of using it to usurp the role and responsibility of the medical profession. Whether it can ever be refined up to a point where society would sanction such a role for a non-human diagnostic system is a matter of speculation that is not within the scope of the present study. Eysenbach (2000) suggests that decision support systems for patients can be useful, but also states that it is not known exactly how consumers use this medical knowledge. This is a challenge for the project.

It is intended to be deployed in a location where members of the public can have easy access. They can then use it in order to help themselves decide to seek help when they have a problem. The location could be a clinic or any other place where members of the public can have easy access to.

An evaluation of its knowledge base was effected. Evaluation of an expert system is vitally important, because it must be verified whether or not the expert system will perform its task satisfactorily. Evaluating the knowledge base of an expert system is part of the evaluation process, and often happens first. If the knowledge base is of broad scope and is accurate, then the program can be expected to be good in its conclusions and hence potentially useful. The evaluation was done as follows. Data on patients with illnesses and symptoms of the study was collected from a primary health care (PHC) clinic. Then FCA implications and the JSM method were applied to the data to generate rules. FCA implications and JSM method hypotheses are equivalent to rules. Hence the term *FCA generated rules*, and the term that FCA is used to *generate rules*. FCA is introduced below. The rules generated using FCA methods were compared to the rules originally given by the experts. The new rules generated by FCA methods were also tested on the field prototype of the CDSS to check if they improved its performance or not. The results were taken back to the experts for discussion, in the iterative manner of the development of expert systems. Because the sample used was of limited size, the results of the study were considered preliminary. Strictly it can be argued that there is no method called FCA implications, but for the sake of the dissertation it has been called a method.

The field prototype of the CDSS which was developed was also used to collect data of the patients' symptoms. This was possible because it uses the same questionnaire that would be used by the researcher. Ethics approval had been granted by the Faculty of Health Sciences, University of Pretoria.

Formal concept analysis (FCA) is a mathematical framework which can be used for deriving causal relationships in data (Ganter & Wille, 1999). FCA methods were used to evaluate the knowledge base of the CDSS which is of the association between certain patient symptoms and illnesses as given by medical experts.

The motivation to use FCA data analysis as well as experts' knowledge is that FCA data analysis may discover some things that the experts may have overlooked. Or at least the experts can review their expertise against actual field data which has been analysed by FCA methods, and perhaps some useful findings can arise. A system like the CDSS cannot be built using FCA data analysis techniques only, involvement of experts is very important. The two FCA methods were chosen so as to compare their results, and it was also thought that they may perhaps complement each other.

The umbrella study area of the project is Consumer Health Informatics (CHI), which is a sub-field of Health Informatics. Health informatics is interested in the development and research of information systems which support medical practice. Health informatics marries computer science with medicine and health practice. CHI is interested in the development and research of information systems which meet the health needs of members of the public. Members of the public are called consumers as opposed to health practitioners for example doctors and nurses. In the dissertation sometimes the term patient is used interchangeably with the term consumer.

Next the areas and themes of the study are listed, and then the study proposition is described.

## 1.2 Areas of Study

- Health Informatics
- Software Engineering.
- Formal Concept Analysis.

## 1.3 Study Proposition

A process (method) for development and improvement of rules of the knowledge base of an expert system using methods of formal concept analysis (FCA) is proposed. The proposed method operates as follows: rules are collected from experts and are implemented in the knowledge base of the expert system being developed. Field data is then collected and then FCA implications and the FCA formulation of the JSM method are applied to the data to generate rules. The FCA formulation of the JSM method will be referred to as simply the *JSM method*. It is so named in honour of nineteenth century English philosopher John Stuart Mill who proposed schemes of inductive reasoning. The rules generated by the FCA methods using the field data are compared to the rules given by the experts. The results are then taken back to the experts for consideration in the iterative manner of knowledge engineering of expert system development.

The method proposed was found to be beneficial in the evaluation and development of the CDSS at the end of the study, though the result was considered preliminary.

To report the study the following approach is used in the dissertation. Firstly the major finding is stated, and then the rest of the dissertation describes how the major finding was reached. In the beginning it was proposed that FCA methods could be used in the evaluation of the rules of a knowledge base of an expert system. An experiment was then carried out. The study design, which follows in the next section, describes the major activities and measures that were done. At the end of the experiment the study proposition was found to be beneficial. A summary of data analysis is made in this chapter to support the findings. The detailed data analysis and the detailed discussion of results, as well as the detailed inner workings of the proposed method follow later, in subsequent chapters. The limitations of the proposed method are also considered in detail in later discussions in subsequent chapters.

## 1.4 Study Design

The aim of the study was to test the rules of the knowledge base of the CDSS against an actual sample of patient data using FCA methods. The results of this could be used to decide the utility of FCA methods at evaluating and improvement of the knowledge base of rules of an expert system. The study design was as follows.

The field prototype of the CDSS program was deployed in a primary health care (PHC) clinic where it was used to collect data on patients' symptoms. A PHC clinic was chosen as the data collection site as it would be easy to find patients with the symptoms and illnesses of the study. Study patients were chosen from among patients who had come for a visit to the clinic. Patient data was collected as follows. The CDSS was used to collect the patient's symptoms using the CDSS questionnaire. It was possible to use it to collect patient symptoms because it used the same questionnaire that would be used by the researcher. The illnesses questionnaire can be found in Appendix II. Thereafter the patient went to see the doctor. Then the treating doctor completed the patient's data by providing the diagnosis. The data for each illness was formatted into a table which contained the symptoms for each patient as well as the treating doctor's diagnosis. There was a table of patient data for each illness. Then the FCA methods were used to generate rules from the patients' data. The rules generated by the FCA methods were compared to rules provided by the experts as follows.

- *Rules Subsumed.* How many experts' rules have been subsumed by the FCA rules? These can be considered to be experts' rules which are accounted for by the FCA generated rules.
- *New Rules.* These are rules generated by the FCA methods which have no corresponding expert rules. These rules may be considered enrichment to the CDSS if the experts' accept them as valid.
- *Rules not Accounted.* These are expert rules which have no corresponding FCA rules as defined above in the first bullet. These rules can be considered misses (omission) by the FCA methods.

In addition, the new rules generated by the FCA methods were tested by measuring the performance of the CDSS at giving advice. By performance the following is meant. How many times did the diagnosis of the treating doctor coincide with that of the CDSS. For example, if out of ten cases it makes the same diagnosis as the treating doctor for seven cases, then its performance will be given as 70%. Its performance was measured twice. First it was measured with the original experts' rules only. Then the new rules generated by FCA methods were

included into the CDSS, and the performance was measured again. The aim was to check whether its performance improved or not, as a result of inclusion of the new FCA generated rules. To do this the data was split into two sets. There was a *training set* which was used for generating the FCA methods' rules. And there was the *test dataset* which was used to test its performance with the new rules included. The results of data analysis were taken back to the experts for consideration in the iterative manner of knowledge engineering of expert systems. The experts had to make the decision as to what should be accepted or not from the FCA analysis.

Consequently its performance was used to give an indication of reliability of its advice. Inspired by a study on a patient medical system called Housecall by Bouhaddou et al (1998), a performance target of 70% was then set for the CDSS.

## 1.5 Data Collection

Patient data was collected at a PHC clinic in Pretoria by the author of the dissertation. Data collection took about 6 weeks to complete. The total number of subject patients collected was 59 patients. The breakdown of study patients according to illness is made in Table 1.1.

Malaria and peptic ulcers' data was not collected for the following reasons. Malaria cases were not found in the study clinic. This is because of the season chosen for the study and because in Gauteng province malaria is not very common. Data collection for peptic ulcers also did not occur because the planned data collection site was in the process of moving to a different location during the period of data collection. So data was collected for only three of the five illnesses which were initially selected, these are: *tuberculosis*, *chronic bronchitis* and *hypertension*. This was accepted because the results of data analysis were found adequate to decide preliminarily the utility of FCA in evaluation of the CDSS. A summary of data analysis is made in the next section. In Chapter 8 the detailed discussion of the results is made.

**Table 1.1 Number of study patients collected.**

Tuberculosis (TB)	31 patients
Chronic Bronchitis (CBr)	7 patients
Hypertension (HP)	21 patients

## 1.6 Summary of Results

**Table 1.2 Summary of comparison of FCA generated rules and experts' rules.**

<b><u>Item</u></b>	<b><u>TB</u></b> <i>8 original experts' rules.</i>	<b><u>CBr</u></b> <i>3 original experts' rules.</i>	<b><u>HP</u></b> <i>No original experts' rules.</i>	<b><u>Total</u></b> <i>11 original experts' rules.</i>
Rules Subsumed	8	1	-	9
Rules Not Accounted	0	2	-	2
New Rules	8	2	4	14

Note on hypertension: At the time of data collection there were no experts' rules for hypertension. However, the decision was made to go ahead and collect data and then apply the FCA methods to see what rules could be generated from the patient data. These were tested against expert opinion.

**Table 1.3 Summary of CDSS performance.**

<b><u>Study Illness</u></b>	<b><u>Performance on Test data</u></b> <i>(original experts' rules)</i>	<b><u>Performance on Test data</u></b> <i>(new rules)</i>
Tuberculosis	80%	90%
Chronic Bronchitis	-	-
Hypertension	-	70%

Note: Hypertension did not have experts' rules when data collection happened, so there are no performance values for original experts' rules. The performance for bronchitis was not done because its dataset was too small to split into two. But it was felt that the available results of performance added meaningfully to the aims of the study as explained in the summary discussion that follows.

The following summary of the result of the study is made at this stage. From Table 1.2 it can be seen that by using the FCA methods 9 out of 11 experts' rules have been accounted for (subsumed). The number of experts' rules not accounted for is 2. This means that the FCA methods subsumed most of what was originally proposed by the experts. In addition, 14 new rules were generated by using the FCA methods. This is 127% more than the original 11 experts' rules. Of course there were no original experts' rules for hypertension, so perhaps the figure would be lower if there were original experts' rules for hypertension. For tuberculosis the numbers of rules generated by the FCA methods were (coincidentally) equal to the number of rules originally suggested by the experts: 8. For bronchitis it was 2. For tuberculosis the performance of the new rules went up when compared to the performance of the original experts' rules, Table 1.3. The experts preliminarily accepted all the new rules suggested by the FCA methods though with some reservations in some instances. For example they observed that the support of some of the FCA rules was low.

The main conclusion made in the study is that preliminarily the proposed evaluation method of a knowledge base using FCA methods may be useful. Perhaps now a full wide scale study with a larger patient sample can be undertaken. The performance of the CDSS was also relatively high, Table 1.3. It was above 70% for tuberculosis and hypertension, this satisfied the target of 70% which had been set initially. This suggests that its advice may be reliable, and hence it may potentially be useful. Again this is a preliminary result.

The detailed discussion of the results and conclusions are done in Chapter 8 and Chapter 9 respectively.

## 1.7 Dissertation Outline

The structure of the dissertation can be summarized as follows:

- Chapter 1 - Introduction (the current chapter)
- Chapter 2 - Consumer health informatics.
- Chapter 3 - Expert systems (Decision support systems).
- Chapter 4 - The design and implementation of the CDSS
- Chapter 5 - Health expert system evaluation
- Chapter 6 - Data collection.
- Chapter 7 - Data analysis techniques.
- Chapter 8 - Results of data analysis.

- Chapter 9 - Conclusion.
- Bibliography and Appendices.

## **Chapter 2: Consumer Health Informatics**

The main considerations in the development of a consumer health informatics (CHI) program are discussed in this chapter. Firstly background is given about the field of CHI which is the umbrella area of study. Following this, the origin of the CDSS project is described. This explains how the project was conceived, the health problem addressed by the CDSS, and how the choices of illnesses for the study were made. This is followed by a discussion of the medical diagnosis process, which discusses strategies for building reliable and accurate diagnosis programs. The conclusion of the CDSS can be loosely considered to be a preliminary diagnosis. In the dissertation it is loosely called a diagnosis. The term diagnosis is not used in the strict medical sense. A review of the ethical and legal issues associated with health information systems is then made.

## **Chapter 3: Expert Systems**

In building the CDSS program, potential examples or styles that could be used have been sought. Artificial Intelligence methods which have been used in decision support systems and diagnostic programs, as well as other methods were reviewed. Expert systems were selected as the design style. In this chapter the main considerations in the development of an expert system are discussed. The chapter begins with background to expert systems. Then knowledge acquisition and knowledge representation for expert systems are considered. Tools that are used in the development of expert systems are also discussed. The chapter ends with some examples of expert systems, some of which are old and well known in the expert system community, and some others which are newer.

## **Chapter 4: Description of the CDSS Application Architecture**

This chapter describes the inner working of the CDSS software program. A description of its role in the health system is given first. Then a description of how the diagnosis strategy of the CDSS was implemented in the software program is made. This includes the description of its main software modules. A discussion of the design of the user interface then follows. This is followed by a description of the project team – the engineer and experts – as well as the roles that they played. Then a description of the evolution of the CDSS software prototype is made. This describes the stages that it would have to undergo before it can be accepted for use in the community.

## **Chapter 5: Evaluation Approach**

This chapter covers the theme of evaluation and improvement of rules of the knowledge base of a health expert system. It begins with an introduction of the evaluation of health expert systems which describes the underlying principles to health expert system evaluation. Two case studies are presented. Firstly is the evaluation of MYCIN which shows, by example, the factors of health expert system evaluation. MYCIN is a classic medical expert system. The second case study is the Predictive Toxicology Challenge (PTC), which is an example of the use of the JSM method. The Predictive Toxicology Challenge is part of the inspiration to use methods of FCA in the CDSS study. The chapter ends with a description of the evaluation approach of the CDSS study.

## **Chapter 6: Data Collection**

This chapter describes the process of data collection. The source of the data is explained, the choice of study site, the sampling strategy, and the formatting of the data.

## **Chapter 7: Data Analysis Techniques**

*FCA implications* and the FCA formulation of the *JSM method* are the data analysis techniques used in the study. The FCA formulation of the JSM method has been referred to simply as the JSM method. The two methods are part of formal concept analysis (FCA). This chapter explains the theoretical foundations of FCA and of the two FCA methods. The following basic definitions of FCA are given: a *formal context*, *formal concept*, the *prime operation*, and a *formal concept lattice*. FCA implications are discussed first, then the JSM method. Automated tools that were used to generate implications and hypotheses are also presented. The chapter ends with some examples of FCA implications and JSM hypotheses.

## **Chapter 8: Data Analysis**

This chapter presents the results of data analysis. The results are presented for the illnesses for which data was collected: tuberculosis, then bronchitis, then hypertension. Malaria and peptic ulcers have not been included due to data collection limitations. For each illness rules generated by FCA methods are listed. These rules were compared to the experts' rules as follows: *rules subsumed*, *new rules* and *rules not accounted*. The new rules were also tested by measuring the performance of the CDSS.

## **Chapter 9: Conclusion**

This chapter summarises the findings of the study, and assesses the extent to which the study proposition had been met. Other findings are also discussed, and then gaps which may have been identified in the study are also raised.

## **Bibliography**

The bibliography comprises the references that were used in the study. There is no ‘literature review’ chapter, which explicitly covers the literature reviewed. Each chapter of the dissertation extensively references the relevant literature sources and their use.

## **Appendix I: Glossary**

Appendix I consists of a glossary of terms which are used in the dissertation. Some of the terms are sourced from the medical and clinical background while others come from the computer science background.

## **Appendix II: Illnesses Questionnaire**

The study has included the ‘illnesses questionnaire’ as an appendix. This questionnaire is the basis of the dialog between the CDSS program and the patient. It is derived from the standard assessment dialog between patient and doctor. This questionnaire was created by one of the experts, the nursing sister, who took part in the study. The sources used included her experience, literature as well as fresh observations of primary health care consultations. The choices of illnesses included in the study are discussed in detail in Chapter 2.

## **Appendix III: Experts’ Rules**

A list of the rules given by the experts for each illness is given here.

## Chapter 2

### Consumer Health Informatics

The main considerations in the development of a consumer health informatics (CHI) program are discussed in this chapter. Firstly background is given about the field of CHI which is the umbrella area of study. Following this, the origin of the CDSS project is described. This explains how the project was conceived, the health problem addressed by the CDSS, and how the choices of illnesses for the study were made. This is followed by a discussion of the medical diagnosis process, which discusses strategies for building reliable and accurate diagnosis programs. The conclusion of the CDSS can be loosely considered to be a preliminary diagnosis. In the dissertation it is loosely called a diagnosis. The term diagnosis is not used in the strict medical sense. A review of the ethical and legal issues associated with health information systems is then made. This includes the ethical profile of the CDSS program. The following therefore are discussed in this chapter:

- 2.1 Introduction to Consumer Health Informatics
- 2.2 Origins of the project
- 2.3 The diagnosis process
- 2.4 Ethical issues
- 2.5 Conclusion

#### 2.1 Introduction to Consumer Health Informatics

The main study area of this project can be classified under the broad subject heading of *health informatics*. Health informatics can be loosely defined as the field concerned with the development and research of information systems which support medical practice. Examples of health informatics endeavours and applications are electronic patient records, clinical decision support systems, medical image recognition software, and development and maintenance of standards for electronic means of health and medical data exchange.

The field of health informatics marries the field of computer science with medicine and health practice. Some of the leading people in the field of health informatics are health practitioners who then studied computer science. For example, Shortliffe is a medical doctor who also obtained a PHD in computer sciences. He was involved in the development of a medical expert

system called MYCIN which has been used as a case study in the project. MYCIN was intended to help doctors with the treatment of infections of the blood. Another leading person in the field is Ball of the Johns Hopkins University in the United States of America, who is a mathematician and computer scientist by background, and who is now a professor in Johns Hopkins University School of Nursing. As is the nature of many research fields today, health informatics is also multidisciplinary. It involves human sciences researchers as well as others from various fields. For example, human science researchers may conduct an evaluation of the social impact of a health informatics intervention.

In South Africa health informatics is a relatively small area of research, but one which may grow substantially in the future, as happened in countries such as the United States of America and the United Kingdom. An example of the growth of health informatics in other countries is that of the government of the United Kingdom which has invested in a multi-billion dollar project called the National Health Services (NHS) Connecting for Health (NHS Connecting for Health, 2006). In the future the same could happen in South Africa.

Within health informatics the CDSS project can be classified under the subheading of Consumer Health Informatics (CHI). CHI is concerned with the development and research of information systems which meet the health needs of members of the public. According to the parlance in this domain, members of the public as opposed to health practitioners are called consumers. An example of a CHI application is the Health24 website based in South Africa (Health24, 2006). NHS Direct Online, in the United Kingdom, is another example (NHS Direct Online, 2006). NHS Direct Online is the web-based wing of NHS Direct, a national telephone and health advice system.

An area covered by CHI systems includes online access for consumers to their electronic health records (Shabo, 2006). The benefit of this could be that each record may contain reminders for the patients, as well as information that has been customized for the particular patients' cases. The online access of patients to their health records has also been proposed as a legal requirement in some countries. Some other consumer health informatics applications involve the self-management of some aspects of some chronic illnesses. Some contain self-care measures needed after an operation (Pallesen, Engberg & Balach, 2006:776).

The measures to involve the consumers more in health practice is driven by greater sophistication of consumers, by cost, and by the value that has been recognized by health care

providers in the involvement of members of the public in health care. For example, in South Africa, the care of terminally ill HIV/AIDS patients is commonly the responsibility of the patients' families. The patients' families need to be equipped for this task, therefore, online resources are potentially valuable.

## **2.2 Origin of the CDSS Project**

This section describes how the research project originated and the motivation for the research project.

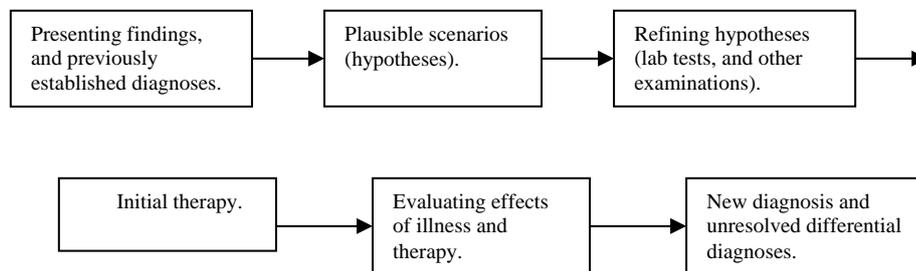
### **2.2.1 How the Research Project Came About**

The author is a researcher at the Medical Research Council (MRC) of South Africa. He was asked to develop the CDSS program. The program is intended to help members of the public (consumers) decide when to seek professional medical help when there is a problem. The program is intended for deployment away from health facilities, for example in computerised community information points designed to provide access to a range of services from government and other sources. The program is intended to operate by prompting the user (consumer) to enter their symptoms via a set of questions which are relayed through a user-computer dialogue. The user-computer dialogue was to consist of simple questions, which are to be answered by selecting one of a defined set of responses, for example, 'yes' or 'no' answers often suffice. The program was to assess the potential that the user was suffering from one of a list of common ailments, such as tuberculosis or hypertension which are problems in South Africa.

After an initial survey, it was suggested that a good way of building this program would be to use an expert system style. The challenge was then to develop and evaluate this program. A field prototype of the CDSS was developed and completed for use in the study. It was found that FCA was applicable in the evaluation and improvement of the knowledge base of the CDSS. Development and evaluation of the knowledge base is often the first task in the development of an expert system.

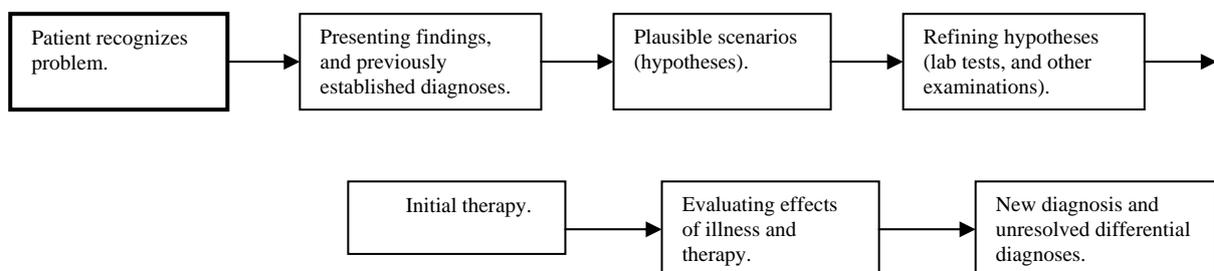
## 2.2.2 Health Problem Addressed

Access to health information is suggested by many research studies as a means to enable patients to be more active participants in the treatment process, leading to better medical outcomes (Eysenbach, 2000). An initial survey done at the beginning of the project suggested that CHI applications that provide such health information have not been well-explored in South Africa. A CHI decision support tool was then proposed to help members of the public decide when to seek help when there was a problem. The program was to focus on symptoms of some common but important (serious) illnesses in South Africa. Below is a description of the role of the CDSS with respect to the health system.



**Figure 2.1 The diagnosis process.**

The diagnosis process can be described using Figure 2.1. The diagram is derived from a description by Miller and Geissbuhler (1999). The role of the CDSS programme is to help member of the public recognize when they have a problem, in order for them to make a timeous visit to the doctor or clinic and be fully assessed and helped. Figure 2.1 can be revised to include awareness of the potential problem by the patient as the first step. This is shown in the Figure 2.2. The other steps are performed by a doctor.



**Figure 2.2 CDSS and the diagnostic process.**

To do data collection, submission of a study protocol to the Department of Health Gauteng province South Africa was done. The following comment was made by the study protocol reviewer of the Department of Health about the CDSS program. He characterized it as, ‘a patient empowerment study that used an innovative way to convey health information.’

### 2.2.3 Criteria for Selecting Study Illnesses

Due to lack of awareness of the early signs of many common illnesses patients tend to procrastinate in seeking professional help when they should. In many regions of the country, the scarcity of medical practitioners and the distance between health consumers and providers exacerbate the problem. Therefore, it was decided that, ideally, illnesses to be selected for the study should be common in South Africa, and should be potentially harmful, but which can be readily treated if detected early (Department of Health, 2003).

**Table 2.1 Illnesses selected for the CDSS** (*Data though was only collected for three; tuberculosis, hypertension and bronchitis*).

<u>Illness</u>	<u>Major Symptom</u>
Tuberculosis	Persistent cough
Bronchitis	Persistent cough
Peptic Ulcers	Pain in the abdomen
Hypertension	Persistent headache
Malaria	Fever

For the sake of the prototype which was developed, and for the sake of the scope of this study, only five illnesses were chosen, Table 2.1. There are other illnesses which meet the criteria of the study which could have been included, such as diabetes. These may be added later in an operational version of the CDSS. Next is a brief description of each study illness.

*Tuberculosis* is an infection of the lungs. If not treated, it can cause death. The experts described tuberculosis and its symptoms as follows: Persistent cough, night sweats (drenching), loss of weight, poor appetite, and sputum production, are its main symptoms.

*Bronchitis* is an illness of the lungs, characterised by a persistent cough which produces sputum that is normally muco-purulent. It can complicate, it is degenerative, and it affects the quality of

life of the sufferer. It is normally associated with cigarette smoking. Some of its main symptoms are persistent cough, heavy tobacco use, sputum production, shortness of breath, chest tightness, wheezing chest.

*Hypertension* is also called high blood pressure. As blood is pumped around the body it exerts pressure on the walls of the arteries. If the pressure is consistently too high then hypertension is diagnosed. If not treated it can complicate, and some of its major complications are stroke, heart attack, kidney damage and loss of vision.

*Peptic Ulcers* are breaks in the lining of the stomach. Most peptic ulcers occur in the first layer of the stomach which protects the stomach from stomach acids which are used in digestion. If an ulcer perforates the lining of the stomach then it is a medical emergency.

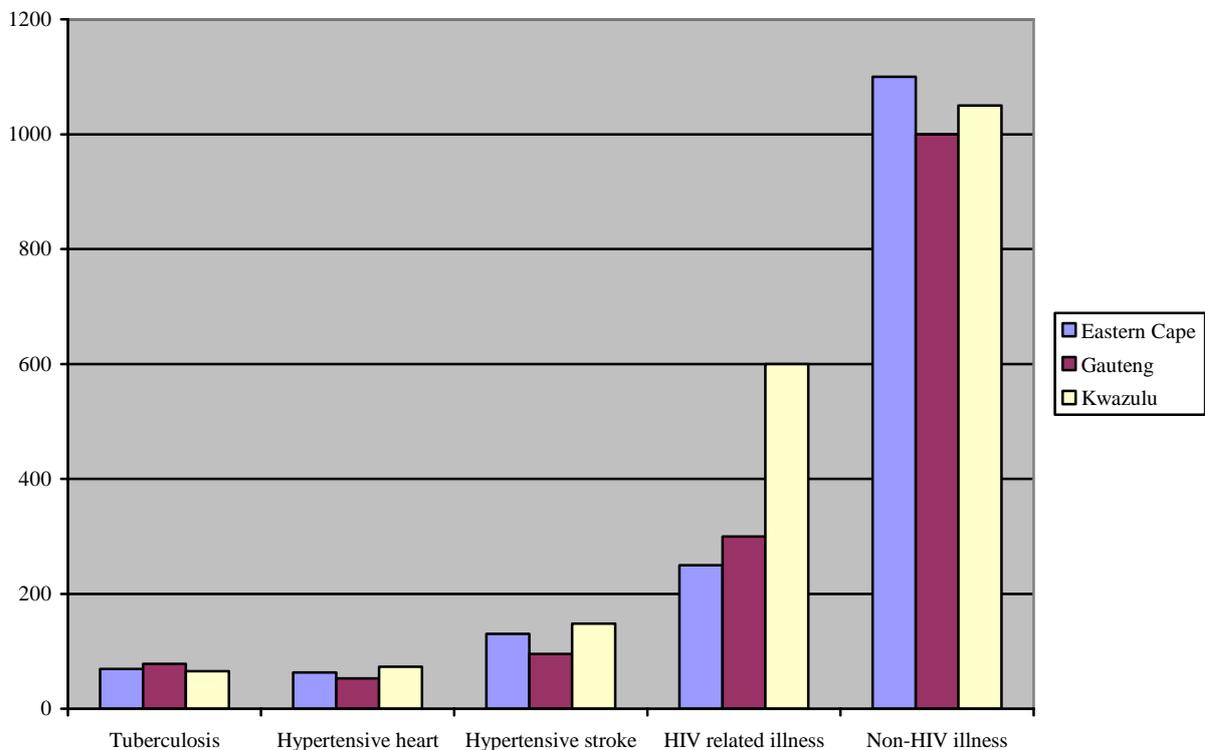
*Malaria* is characterized by a fever which is caused by a parasite that is spread by malaria bearing mosquito. It is often described as a tropical disease. If not treated it can complicate, and even cause death.

**Table 2.2 Summary of burden of disease index (for three provinces)**

<u>Province</u>	<u>Tuberculosis</u>	<u>Hypertensive heart</u>	<u>Hypertensive stroke</u>	<u>HIV related illness</u>	<u>Non-HIV illness</u>
Eastern Cape	69	63	130	250	1100
Gauteng	78	53	95	300	1000
Kwazulu	65	73	148	600	1050

*(Mortality rates are per 100 000 persons, the mortality for tuberculosis and hypertensive disease are non-HIV related, they were taken for the year 2000)*

To support the choices of illnesses, a table summarising mortality for non-HIV related tuberculosis and hypertension in South Africa has been created by the author, Table 2.3. Stroke and hypertensive heart are related to hypertension. The table has been created from the Medical Research Council of South Africa's burden of disease study summary report (Bradshaw et al, 2000). Three provinces are used in the table. Data for chronic bronchitis, peptic ulcers and malaria were not available in the summary report. But from the discussions with the medical experts who took part in the study it was agreed that they were a problem.



**Figure 2.3 Burden of disease index.**

From the table the chart of Figure 2.3 was created. It can be seen clearly that two of the chosen illnesses, namely tuberculosis and hypertension are a problem, constituting about twenty to thirty percent of non-HIV related deaths. In future studies it would be interesting to see to see whether models can be used to measure the impact of a program like the CDSS on these problems.

## 2.3 The Medical Diagnosis Process

### 2.3.1 Background

A process similar to that employed by a doctor to assess a patient is employed by the CDSS program. This approach is similar to how a consumer health system such as *Housecall* was derived from an existing system called *Illiad* which was meant to assist doctors (Bouhaddou, Lambert & Miller, 1998). The doctor may have used *Illiad* when diagnosing a difficult case. For the consumer *Housecall* can be considered a sophisticated medical reference. As a bit of

context an explanation is given of *Illiad* and *Housecall*. *Illiad* was an expert system meant to help doctors in diagnosing diseases of internal medicine. It had over one thousand five hundred (1 500) diagnoses in its knowledge base. Its combined count of illnesses and symptoms is over six thousand (6 000), it includes some that are common and others that are rare. The intention of *Housecall* was to use the knowledge base of *Illiad* in order to develop a patient medical application.

How patients digest and act on medical knowledge is not well known. However, a study on the use of the internet by patients could give a clue (Schwartz et al, 2006). The study by Schwartz et al found that a high percentage of patients had used the internet to seek health information, before going to the doctor. A major difference between *Illiad* and *Housecall* was that the conclusion ('diagnosis') of *Housecall* were always less than seventy percent certain while *Illiad* was higher. The doctor could enter lab tests and other examinations into *Illiad*, while *Housecall* was based on patient symptoms only.

A study on *Housecall* in the USA had some of the following results. *Housecall* was used on average once per month by each user. *Housecall* was used for the following reasons: to obtain a second opinion, to help with questions to ask the doctor and to practice self-care. Under self-care the wish was that users wanted to avoid continuous visits to the doctor for minor ailments, for example, a sore throat. It was also used to help family and friends. The average user of *Housecall* was relatively well educated (college graduate) with a high salary, on average fifty thousand US Dollars per year. *Housecall* was a commercial system that users had to buy.

The target audience of the CDSS program though is expected to be poorly educated and to fall in the low income groups in South Africa. Such people are also concerned about their health. It is suggested that if they are given means to be more proactive in their health, they will respond positively (Carty & Kenney 2006:36).

The following discussion describes the medical diagnosis process, and presents some strategies for modeling medical diagnosis.

### **2.3.2 Modeling Diagnostic Reasoning**

In 'Heuristic Methods for Imposing Structure on Ill-Structured Problems', Pople (1982:6) suggests that the greatest effort (the largest effort and time) in problem solving involves the

formulation of the problem. He suggests that some problems are ill-structured problems (ISP). If one is to successfully solve an ISP, then one has to be able to formulate it as a well-structured problem (WSP). Thus, the greatest problem solving effort goes into transforming an ISP into a WSP. He suggests that with an ISP the options are not necessarily known in advance, but are generated and refined as more data is acquired about the problem. A major part of the problem solving strategy would be first to come up with very general options and constraints, which could help gather information. From there, as more information is gathered, it is possible to become more specific and to focus on a problem. The other part of the strategy is to break down the problem into smaller more manageable chunks. He further suggests that the problem solving task involves a large search process. Required information may often show up late and sometimes as if by accident. He adds that the problem begins to gain structure as it is 'digested' into manageable chunks (components) and by evoking many different requirements in testing the various options.

He then describes two approaches to structuring the medical diagnosis problem. The first approach is binary, the options being yes or no. For example, for each symptom (presenting finding) a possible cause (or causes) is assigned. The program then tries to prove if the possible cause hypothesized is true by invoking information. If it is found to be false, then the next option is explored and so on. To structure the search, the obvious or common causes are explored first. For example, if someone presents flu-like symptoms in winter, it is natural to explore the possibility of flu. Epidemiological information is used in the diagnosis process in this manner. Another strategy is to consider groups of findings which are known to be common to an illness and to use these to generate hypotheses, which are then also tested. The hypotheses in both strategies are helpful if they are mutually exclusive. Nosological classification can help make this possible. Nosology is the study of the classification of illnesses. An obvious classification strategy is by organ. For example, disease of the heart can be a category, diseases of the alimentary canal can be another category. Nosological classification was used by Aikins (1983). To illustrate the notion of mutual exclusion, an example from Pople is considered. If someone has a fever, it could be an infection, some form of cancer, a cardiovascular disorder and so on. But if the doctor were to inquire whether the person had experienced shaking chills and the answer was affirmative, then the options would now be limited to infections and the other options would be excluded. Mutual exclusion allows only one to be true among a set of options. Generating the mutually exclusive sets is another problem of course.

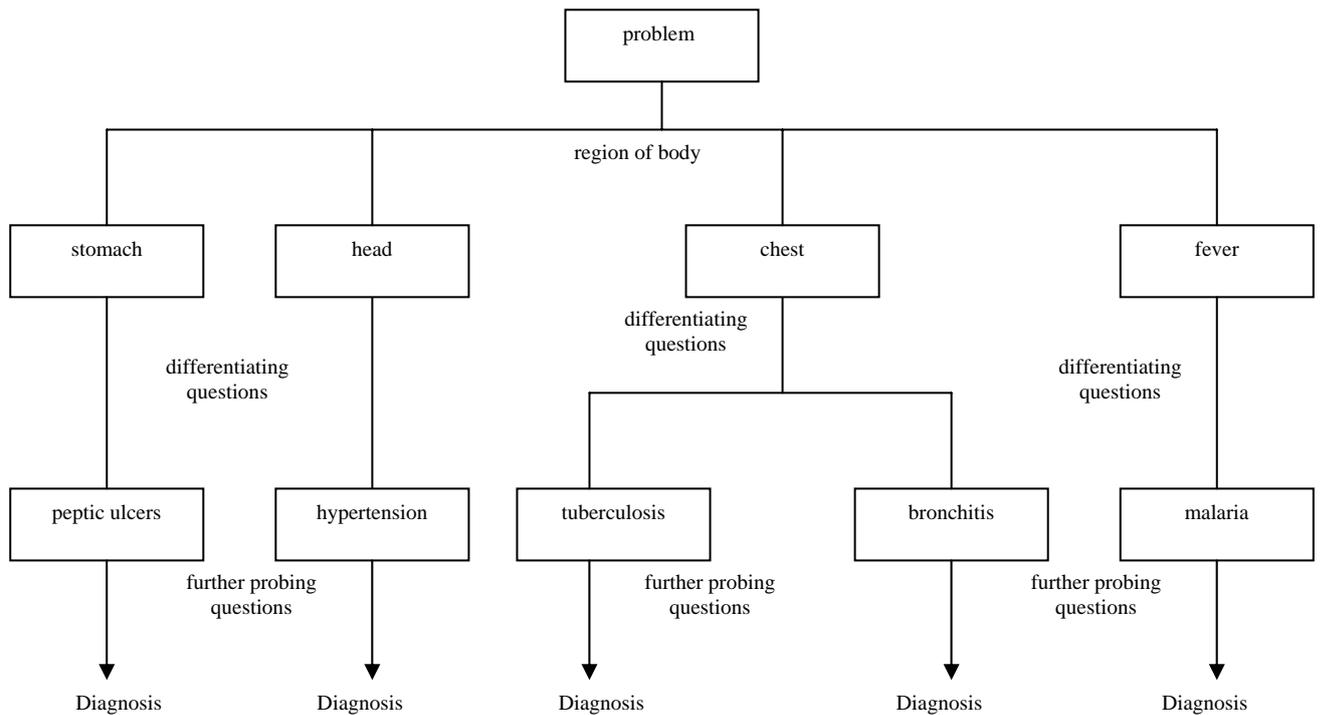
Pople (1982:9) suggests that during practice a doctor may need an external reference to aid in diagnosis when there is an unusual presentation of symptoms, or if it is the case of a rare condition. In most cases a doctor is able to recollect previous cases to use as a solution. Protocols and procedures have also been developed for dealing with well defined clinical problems. The Standard Treatments Guidelines and Essential Drugs List of the Department of Health is an example of this (Department of Health, 2003).

To demonstrate, Pople created Internist which was a decision support system meant to assist doctors with the diagnosis of illnesses of internal medicine. Its conclusions were weighted lists of differential diagnoses. They were weighted according to the probability that they accounted for the patient's symptoms. When Internist was tested, in most cases the actual diagnoses of test patients was on its conclusion. This was in over 90% of the test cases.

### **2.3.3 CDSS Diagnosis Strategy**

The CDSS diagnosis strategy is given by the following steps Figure 2.4 is a summary of the strategy.

- Step 1. The CDSS enquires about the area of the body in which the problems presents. For example, chest is the area for tuberculosis and bronchitis, while stomach is the area for peptic ulcers. If the patient fails to identify the body part of the problem then the CDSS suggests a fever by asking, in which case malaria is the possibility.
  
- Step 2. The CDSS checks for certain key symptoms that differentiate illnesses in the part of the body that they occur, in order to make a tentative diagnosis. For example tuberculosis and bronchitis are differentiated by weight loss and cigarette smoking. Weight loss would apply more to tuberculosis, while cigarette smoking applies more to bronchitis. Both illnesses are characterized by a persistent cough. If a patient has a persistent cough and has lost weight then tuberculosis is suspected tentatively.



**Figure 2.4 CDSS diagnosis strategy.**

- Step 3. It asks further probing questions which are in line with the standard symptoms for the illness tentatively suggested. This is in order to confirm presence of the illness. For example with tuberculosis it will ask if the person has an appetite, if they experience extreme night sweat, or if they have been in contact with someone with tuberculosis. The full lists of standard questions for each illness are given in Appendix II.
- Step 4. At the end of the questions it uses the combination of rules given by the experts to make a diagnosis. For example for tuberculosis the following is a combination {PC WL NS}. It states that if a patient has persistent cough (PC), weight loss (WL) and extreme night sweat (NS) then they may be suspected to have tuberculosis. All the combinations of rules for the illness are given in Appendix III.

Ideally the CDSS should offer the following advice options. When it is able to conclude with an illness, it should name the illness and then advise the user to seek help at a PHC clinic or doctor. If it fails to diagnose, but considers the user to have serious symptoms then it should suggest that the user go to the clinic for further assessment, and also should list two or three illnesses that may be the cause according to its rules. If the user has non-serious symptoms, then the user may not have a serious illness. In the last case there is an ethical challenge. What if it fails to identify

a serious illness, but it is later found that the person had a problem. Then it needs to be evaluated to see what possible harmful outcomes can occur if it finds no problem yet there was actually one. Another strategy would be for the CDSS to guarantee finding a problem. For example if it were to find all cases of patients with problems, but also have the error (short coming) to say there is a problem when there is actually none for many cases. This is acceptable when compared to saying there is no problem when in fact there is a problem, even if these cases are few when compared to correctly diagnosed cases. But a strategy that guarantees identification of the problem always cannot be found (Miller & Geissbuhler, 1999). A possible answer to this problem is extensive testing and evaluation to identify potentially harmful outcomes when it errs. Its benefits to help, also have to be weighed against its disadvantages. In the current study it refers all users to the clinic regardless whether it considers their problem serious or not. Under this approach it is assumed that most people who come to the CDSS for help do actually have a problem, but have not yet decided to seek help. This was agreed upon by the CDSS development team which included medical experts. Those who do not really have a problem but who come for help are assumed to be small in number compared to those with a problem. In the current study the focus of evaluation is on rules which make positive predictions. Rules which make negative predictions, that is, those which say that there is no serious illness have not been included.

The conclusions of the CDSS are either positive for an illness, or it does not make a conclusion, in both cases it refers the patient as explained above. When it does not make a conclusion it could be that the patient has another illness. Or the patient does not have a serious illness. Or it could be that it has failed to find the illness though it was present. But it was still considered to be useful for encouraging patients to seek help. Later it may be refined to address these other options.

## **2.4 Ethical Issues of Decision Support**

Some work on ethical issues of clinical decision support systems has been done, a summary of which follows (Goodman, 1999:217). The discussion is broad, in order to give background. It concludes with a review of the ethical profile of the CDSS program.

### **2.4.1 Standard View of Diagnosis**

In early work on clinical information systems the ‘Standard View of Diagnosis’ was developed (Goodman, 1999:218). It is on the side of caution. The reference interprets it as saying that in principle, or at least in the foreseeable future, computers cannot be used to supplant human decision makers, or that a computer diagnosis cannot be used to override a human decision or diagnosis.

In the discussion of ethical issues three concerns are suggested: care standards, appropriate use and appropriate users, and professional relationships.

### **2.4.2 Care Standards**

The main question to be asked here is: Does the new system improve patient care? If the answer is no, or is not clear, then ethically there is a problem. A follow-up question is then asked: Does the new system impose any new or special responsibility? A major responsibility of care standards is to cause no harm. The question to be asked: What harm can result from use of the system?

### **2.4.3 Appropriate Use and Users**

A tool can be abused in the following three main ways. Firstly it can be used for purposes for which it was not intended. Secondly it can be used without adequate expertise or training. And thirdly it can be used incorrectly, for example carelessly or sloppily.

The first instance could be to use a tool meant for modest decision support in a way as to abandon professional advice. A scenario can also be imagined where an insurance company or managed care organisation uses a computerized decision support tool to evaluate or challenge clinical decisions. This would also be an example of abuse.

### **2.4.4 Professional Relationship**

It is suggested that there should be ‘shared decision making’ in the treatment process, between doctor and patient. It needs to be checked how the tool interferes or helps this relationship.

## 2.4.5 Legislation of Decision Support Software

The FDA in the United States of America regulates various health related decision support software. The following are the main categories considered by the FDA.

- *Educational and bibliographic software* is not regulated.
- *Software components* incorporated into medical devices are regulated, for example, software used in pacemakers, clinical laboratory instruments, and ventilators.
- *Software accessories* which are used in connection with medical devices are also regulated. Examples are radiation treatment planning, and calculation of rate response for cardiac pacemakers.
- Stand alone software is the ‘controversial class’. The decision to regulate or not has not been made. Some examples of this kind of software are:
  - Blood bank software systems that control donor deferrals and release of blood products.
  - Software for helping practitioners to arrive at diagnoses.
  - Software which analyses potential therapeutic interventions for a particular patient.

## 2.4.6 Problem of Illiteracy

It is suggested that there is a high level of illiteracy in South Africa. It can then be asked if a tool like the CDSS is ethical if it can only be used by less than forty percent of the population of the country?

## 2.4.7 Ethical Profile of CDSS

An application for an ethics review of the CDSS program was made to the Ethics Committee of the Health Sciences faculty at the University of Pretoria. There were no ethical objections raised and the study was approved. In addition to the ethics committee approval the following potential ethics issues were identified which may have to be addressed before it may be deployed for wide scale public use.

There are many consumer decision information aids on the internet. Many are sophisticated, such as *MedlinePlus*, a consumer medical encyclopedia operated by the United States of America National Library of Medicine. Although there is a disclaimer that the knowledge should not be used for treatment or diagnosis of medical conditions its content is very strong

covering many common as well as rare illnesses. The important question is how consumers use that information. Eysenbach (2000) suggests that this is not well known. The ethical challenge that has been identified in this project is how a potentially unsophisticated user like the ones envisaged in South Africa may use the CDSS. Would they see it as a substitute for a doctor's assessment? Because this is its intention, a study needs to be conducted to determine this.

The second concern is safety. The question is what potentially harmful outcomes can arise from using it. Another study and one that is wide ranging in terms of sample size and measure needs to be conducted before it can be deployed. The results of the current study, though promising, were considered preliminary.

The effort to develop the CDSS program is justified considering that there is a movement that supports greater access to health information for the health consumer around the world (Wong, 2006:9). In South Africa, the conditions which have been included have caused a great number of deaths and harm, due to a lack of awareness by members of the public.

## **2.5 Conclusion**

An initial survey at the commencement of the study found that CHI is a relatively new field in South Africa and that there were few CHI applications dedicated to health problems encountered in South Africa. The CDSS program was then proposed to address information needs for some illnesses found in South Africa. The illnesses that were chosen for the study are also a priority of the department of health.

In this chapter the background to the project and the reasons why it is necessary to consider a system like the CDSS in South Africa were given. This discussion included the health problem addressed by the CDSS and its role. Then a discussion on its diagnosis strategy was done. This discussion used literature references and other examples. In the discussion its limitations at giving advice were noted, and how this affected its design. The experts' rules used in its knowledge base were shown to be vitally important and hence the evaluation and development method which has been proposed by the study. Next expert systems are discussed, which is the style used in its development.

## Chapter 3

### Expert Systems

In building the CDSS program, potential examples or styles that could be used have been sought. Artificial Intelligence methods which have been used in decision support systems and diagnostic programs, as well as other methods were reviewed. Expert systems were selected as the design style. In this chapter the main considerations in the development of an expert system are discussed. The chapter begins with background to expert systems. Then knowledge acquisition and knowledge representation for expert systems are considered. Tools that are used in the development of expert systems are also discussed. The chapter ends with some examples of expert systems, some of which are old and well known in the expert system community, and some others which are newer. The chapter has the following subheading.

- 3.1 Introduction to expert systems
- 3.2 Knowledge acquisition
- 3.3 Knowledge representation
- 3.4 Explanation
- 3.5 Tools of expert system development
- 3.6 Examples of expert systems
- 3.7 Conclusion

### 3.1 Introduction to Expert Systems

#### 3.1.1 Definition of an Expert System

The main features of an expert system are captured in the following broad definition: an expert system is a program which reasons using a store of knowledge derived from experts and is used for problem solving. This knowledge store is called a *knowledge base*.

‘...An expert system may completely fill the role of a human expert, or it may play the role of an assistant to a human decision maker...’ (Jackson, 1999:2). An expert system can be used to assist a human collaborator attain expert levels of performance. The human collaborator may be an expert, in which case they may be using the program to improve productivity. The allocation

of function between person and machine often determines the success of an expert system. Between the two collaboration schemes suggested above, the CDSS program intends to improve the ability of a member of the public to assess their own health problem and know when to seek professional help. The member of the public using the CDSS program does not necessarily attain expert level performance as a result of using the program, but their ability is improved greatly.

The methods used for problem solving by expert systems are described as heuristic rather than mathematical and algorithmic. Expert systems are used often in decision support, hence the choice for the CDSS.

### **3.1.2 History of Expert Systems**

Expert systems were initially studied as part of the field of *Artificial Intelligence* (AI). AI is a field of study interested in the emulation of human behaviour by machines.

The early period of AI research is called the classical period, this begins after World War II, up to the mid-nineteen seventies. Classical AI was interested in designing programs that simulated or extended human mental capabilities. It was a multidisciplinary field spanning computer science, psychology, philosophy, and linguistics. Research into cognitive science at this time led to the proposal of many techniques of knowledge representation which were very useful for computer science. The knowledge representation schemes that were considered in this period include: *production rules*, *network systems* and *logical formulas*.

Jackson (1999:27) talks of the modern period of AI from the late nineteen seventies to the late nineteen eighties. In this period there was research into AI which argued that the power of the problem solver was not in a sophisticated inference scheme but was rather in the explicit representation of knowledge. Expert systems arose out of this point of view, and MYCIN is an early example. The modern period ends with less optimism, Jackson calls it the 'AI winter'. Expert systems and AI had not delivered on promises that were made. He goes on to blame the failure on over-marketing by AI companies who had built software packages that were poorly integrated with other commercial systems and which ran on specialised hardware. Jackson ends by saying that perhaps there were too many expectations and that AI was not a 'magic bullet' and that not all domains were suitable for AI.

Jackson adds that the advent of the World Wide Web spawned an ‘AI spring’ and that this is as a result of the expanding information environment. He asserts that AI perhaps may have something to offer and that it should go ‘mainstream’.

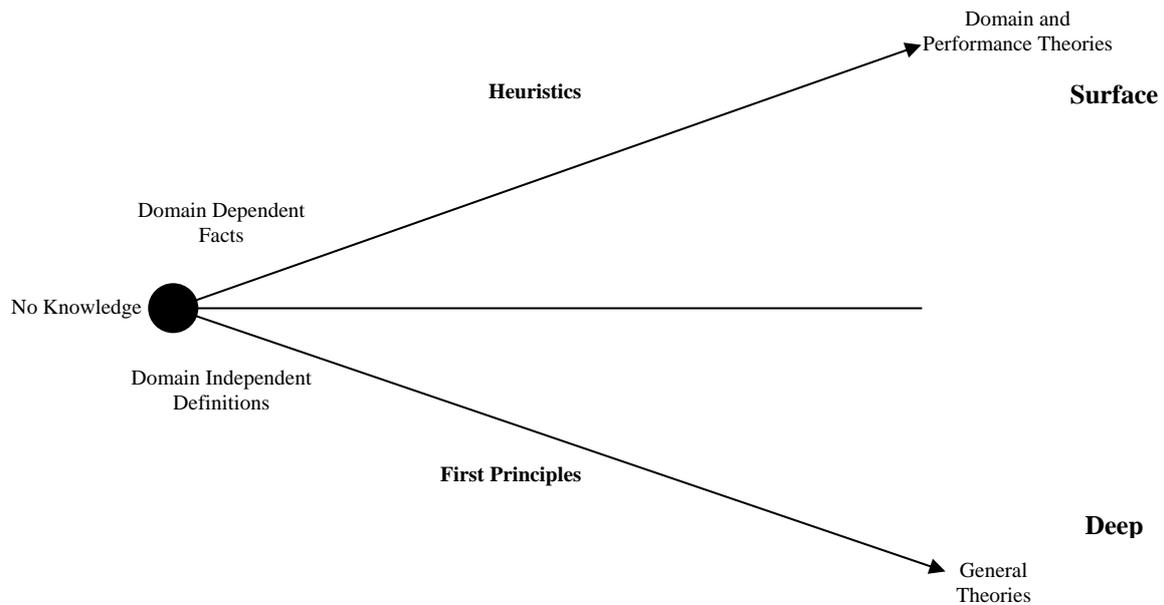
In the project the use of expert systems is viewed as a software engineering endeavor rather than an AI endeavor. Expert systems represent a design style which was found to be applicable to the problem of developing a decision support system to assist members of the public with health problems.

Next some key concepts in the development of expert systems are discussed. The first is *knowledge acquisition*, which explains how the knowledge is extracted from the experts. Secondly *knowledge representation*, which is how the knowledge from the experts is represented in the program. Then lastly *explanation*, which is how the expert system justifies its conclusion to the user. For example a doctor using a diagnostic program may want to know how the program arrived at its conclusion.

## 3.2 Knowledge Acquisition

Jackson (1999) states that knowledge acquisition ‘...is the transfer of knowledge from an expert (knowledge source) to a program...’. It is usually accomplished by interview of the expert by the knowledge engineer. It is further suggested that this process can be accomplished at a rate of about two to five units of knowledge per day. A single rule can be considered to be a unit of knowledge. The following may be the reasons why this may be a slow process. Firstly it is not easy to get an expert to communicate his/her knowledge in everyday terms and the concepts behind the terms need to be analysed and understood by the knowledge engineer. Secondly concepts of a domain often cannot be characterised precisely in a deterministic manner. And thirdly the knowledge of the expert is combined with a lot of common sense knowledge about the everyday world which can be very broad, and is often subsumed, assumed, and as a result it may be difficult to articulate.

Some researchers have considered *automated knowledge elicitation*, while others have considered *machine learning*. With automated knowledge acquisition an expert is interviewed by a machine, and through a person-machine dialogue, knowledge is gathered. In machine learning a program is given a set of examples from which it can identify concepts. The concepts could be rules or any other knowledge representation scheme.



**Figure 3.1 Knowledge model of an expert system (Harmon and King, 1985:33)**

Figure 3.1 is a model which suggests the location of heuristic knowledge with respect to other knowledge of a domain. Heuristic knowledge is favoured by expert systems.

### 3.3 Knowledge Representation

Knowledge representation refers to the formal representation of knowledge of a domain in a program so that it can be manipulated symbolically. Symbolic manipulation is an AI concept which focuses on manipulation of concepts which are represented by symbols (strings) in a program. The scheme is described by *syntax* and *semantics*. The syntax describes the form of the representation language, and semantics is the meaning of expressions of the language. CLIPS is a rule-based knowledge representation language which was used in the project. A good knowledge representation language should also have the following qualities of programming languages: it should be expressive and powerful, simple and convenient, and should facilitate communication (Sebesta, 1999:8).

The three main knowledge representation schemes generally employed by expert systems that have been considered are: *production rules*, *networks* and *objects*.

### 3.3.1 Production Rules

A production rule has a premise and a conclusion. The premise is a set of conditions under which the rule will be activated. When a rule is activated, in an expert system, a new fact may be inserted into working memory or it may execute an action. An example of a rule in the CDSS program is: *Weight loss, Night sweat* → *Tuberculosis*. This rule states that if a person experiences weight loss and extreme night sweats then tuberculosis may be suspected. Rules originate from the intuition that humans solve problems by applying a set of rules of the form that have been described.

The greatest value of rules that was observed in the CDSS project was that they facilitated communication between experts and the knowledge engineer. It is suggested that when interviewed, experts often will not tell the exact whole sequence of how they solve a problem, as this may be something in-built. Applying cases scenarios and asking for potential rules arising was found to be beneficial. For example, the following may be asked of a health expert; ‘...if someone is only experiencing dizziness would you consider that hypertension...’. They may answer by saying that if the patient also has someone in their family who has hypertension then it may be suspected that the patient may have hypertension and that a blood pressure test must be done.

Rules also allow inference schemes (problem solving schemes) like forward and backward chaining (Jackson, 1999).

A rule-based system works by matching information (facts) provided by the user against rules in the knowledge base. Those rules whose premises are matched by facts provided by the user, are gathered and one of them is fired according to the conflict resolution scheme. An example in the CDSS program is given. If, during the interaction, the user answered yes to weight loss and to extreme night sweat, then the rule: *Weight loss, Night sweat* → *Tuberculosis* will activate. Working memory of the program will then be updated with a new fact: Tuberculosis. This new fact (new information about a patient case that has been inferred from other information) may trigger another rule to activate. However, in the above instance the CDSS would conclude that the probable diagnosis for this patient is “Tuberculosis”. In a rule-based expert system working memory is compared to the knowledge base. Working memory is the space used by an expert system when processing a user.

Rule-based systems also have disadvantages. They often contain many *control rules* which are not really concerned with the problem domain, but are to do with the control of the execution of the program (Aikins, 1983). The conflict resolution strategy may also be difficult for a large rule set, as well as uncertainty handling.

Many health expert systems for diagnosis, including MYCIN, are rule-based and for many, their performance has been found to be good. A recent example of a ruled based expert system is EAGLE (Project EAGLE, 2006). EAGLE is a legal advice program developed in England. Its performance was also found to be good. The many examples of successful rule-based expert systems were one of the main reasons why a rule-based expert system was used for the CDSS program.

### 3.3.2 Networks

A network consists of a system of nodes and links. The nodes of a network are concepts or elements of the domain and the links describe the relationship between the nodes. Examples of networks are trees, nets and frames. The network can be hierarchical, therefore, it may be argued that objects are an example of a hierarchical network.

The question to ask of networks is what do the nodes represent, and how can information processing and reasoning be done efficiently. In the CDSS program a decision tree has been used to model the diagnosis strategy, Section 2.2.3 Chapter 2. Once the patient has answered the required questions the program employs the concluding rules.

There are some clinical diagnosis programs whose primary knowledge representation structures are networks, for example INTERNIST (Pople, 1982:1).

### 3.3.3 Objects

Object-orientation (OO) is an approach to software development whose central notion is an object. An object encapsulates data and functions. The functions define behaviour of the object and operate on the data. Objects correspond to the philosophic view that things in life have behaviour (or actions), and attributes (data). For example, a human object has attributes such as height, weight, language, and country or province of residence. A human may have behaviour such as walking, working, eating and thinking. Objects in an OO model correspond to real world

objects. Therefore it can be said that they directly support the intention of a computer program to model real world processes. For example, the CDSS program is intended to model the health assessment process for members of the public. OO supports a hierarchical structuring of objects that reflects “inheritance hierarchies” in the real world. For example, since human beings are mammals, they inherit the features of mammal in general. Hence, in some OO system, mammals and humans could be represented as objects in such a way that human objects automatically inherit features of mammal objects. OO has been used in the development of the control module of the CDSS program.

### **3.4 Explanation**

Expert systems traditionally have a facility to explain program behaviour and recommendations. Such an explanation is typically nothing more than a statement of the sequence of rules that the system has used to arrive at a conclusion. This is important for the following reason: users, knowledge engineers and domain experts need to satisfy themselves that the program’s conclusions are correct and that the knowledge is applied correctly.

Explanation is also linked to evaluation, which is an important part of the development of an expert system. It is often necessary for an expert system to have this facility in order for an expert evaluator to assess general performance.

In the CDSS program an explanation is made to the user after an interaction.

### **3.5 Expert System Development Languages**

Production rules were chosen for development of the knowledge base, and the rest of the CDSS was developed using Java which is a procedural language. This section covers the development languages that were considered, these are: *production rules*, LISP, PROLOG and *procedural languages*.

#### **3.5.1 CLIPS**

CLIPS is a production rule language which has been used in CDSS. Production rule languages have syntax for encoding rules, as well as facilities for managing a knowledge base of rules.

Production rules though may lack some of the features of standard procedural languages like recursion found in C++, for example. Their integration into other systems may also be difficult. It is suggested that production rule languages should be combined with other standard programming paradigms in order to balance their shortcomings while maintaining their strength as production rule languages. In this project an API called JESS which is a Java based encapsulation of CLIPS has been employed. It allows interoperability and data exchange between a rule-based system (CLIPS) and a Java program. From within a Java program it was possible to make calls on the CLIPS rule-based system and to exchange data between the two. This was found to be beneficial.

### **3.5.2 LISP**

Symbolic programming is a concept derived from AI. LISP is the language originally used in symbolic programming, and it can be considered to be a general purpose programming language. It was used in building some of the early expert systems, but later, production system languages were created which were designed for building expert systems. When using LISP to develop an expert systems a rule interpreter for example must still be built, which is something that is in-built with CLIPS.

### **3.5.3 PROLOG**

PROLOG is a logic programming language that was considered for the CDSS. The criticism that was made was that the formalism of rules had to be mapped into the logic system, which would have been a considerable level of indirection in the case of CDSS. Many successful expert systems though have been developed using PROLOG.

### **3.5.4 Procedural Languages**

Some examples of procedural languages are C++, Java and PHP. These languages can be used to represent rules, but the elegance of a production system language is lost when they are used instead. When used to build a rule-based system, a rule interpreter has to be developed and other features which are required in a rule-based system. These features often come standard with a production system language. But the benefits of procedural languages for development of expert systems have been given by the example of JESS. JESS is a Java API and which allows

interoperability between CLIPS, a rule-based programming environment, and Java, a procedural language.

## 3.6 Examples of Expert Systems

### 3.6.1 MYCIN a Classic Expert System

MYCIN is a rule-based expert system developed at Stanford University by the Stanford Heuristic Programming Project (SHPP), in the nineteen seventies. SHPP built DENDRAL which is credited as the first AI program that employed a large body of knowledge to solve a problem. The DENDRAL project was begun in 1965. DENDRAL was a program for identifying chemical structures from mass spectrograms. It used production rules, but it also employed a ‘heuristically shaped combinatorial search space’ (Newell, 1984:xiii), which is a strategy of classic AI problem solving. MYCIN used a very simple limited search strategy which is non-generative. The search strategy was not the main feature of MYCIN. Newell (1984:xiii) also states that MYCIN could be viewed ‘...purely as a body of knowledge...’. The MYCIN project developed what are now common features of expert systems such that it is often called the ‘granddaddy’ of expert systems. In its development it also used many features and lessons learnt from the development of DENDRAL. MYCIN began from collaboration at Stanford University, between Cohen (Chief of Pharmacology), Axline (Chief of Infectious Disease Division), Buchanan (a senior researcher of SHPP), and Shortliffe, who was then a medical student doing a computer science course. Others were quickly attracted to the project as it was perceived as very exciting and interesting. The collaboration began in 1971 and received its first major research grant in 1973.

MYCIN’s domain of operation is the treatment of blood infections. The aim of MYCIN was to help physicians make the most appropriate treatment for an agent which had infected the blood of a patient. Some drugs are effective against certain agents, while against others they are not effective. Some which are effective are very toxic. For example, they are used in laboratory experiments where it is safe or in chemotherapy where desperate measures can be taken. So the role of choosing a treatment for blood infection is not trivial, and a system of this kind could be helpful. The evaluation of MYCIN is a case study discussed in Chapter 5.

### 3.6.2 Recent Examples

Although expert systems may have lost the early prominence of the ‘modern period’ of AI there is still interest. This section on recent examples of expert systems has been split into two categories. First are research examples, and then commercial examples.

#### 3.6.2.1 Research

Sailer, of Ulm University Hospital in Germany, has developed a case based decision support system meant to help novice nurses in the decision making process (Sailer, 2006). He presented a poster at the 2006 International Orem Society conference on nursing practice in Johannesburg.

A system called *EAGLE* was developed in England. Its domain is legal advice. It was intended to help generalist advisors in the offices of the Citizens Advice Bureau. It employed a rule base together with a case database. On evaluation the system showed good performance. Information on the project can be found at <http://www.projecteagle.org.uk>.

#### 3.6.2.2 Commercial

*RuleBurst* is a software organisation which specialises in software for complex business rules, governance rules, and risk and compliance management. It has methodologies as well as processes in its services suit. It was involved in project *EAGLE*. Information on *RuleBurst* can be found at <http://www.myiqms.com>.

*KANA* provides knowledge management solutions. One of its products is *KANA IQ* a tool with case-based reasoning ability that can be used for diagnostic-like tasks. It is intended to help in call centers which have large databases which need to be searched when solving a clients’ problems. Information on *KANA* software can be found at <http://www.kana.com>.

### 3.7 Conclusion

Expert systems were used in the development of the CDSS. This was done because it was found that many decision support systems, of which the CDSS is an example, were built using the

expert system style. Some of these systems are old like MYCIN while others are newer, and many of them were shown to have good performance. There were also many failures of expert systems. Chapter 4 follows next on how the CDSS was actually developed. Then Chapter 5 on evaluation of health expert systems.

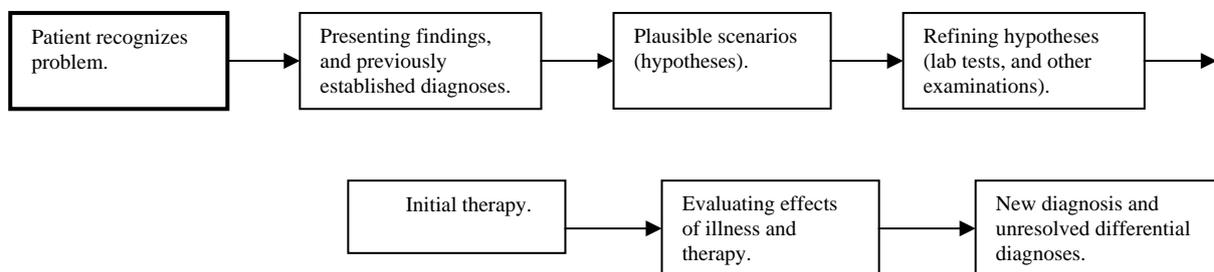
## Chapter 4

### The Design and Implementation of the CDSS

This chapter describes the inner working of the CDSS software program. A description of its role in the health system is given first. This description is a summary of the discussion which was made in Chapter 2, it is restated here in order to give context to design and implementation of the CDSS. Then a description of how the diagnosis strategy of the CDSS was implemented in the software program is made. This includes the description of its main software modules. A discussion of the design of the user interface then follows. This is followed by a description of the project team – the engineer and experts – as well as the roles that they played. Then a description of the evolution of the CDSS software prototype is made. This describes the stages that it would have to undergo before it can be accepted for use in the community. The chapter uses the following subheadings:

- 4.1 Role of the CDSS
- 4.2 Software implementation of the CDSS
- 4.3 User interface design
- 4.4 Project team
- 4.5 Project evolution
- 4.6 Conclusion

#### 4.1 Role of CDSS



**Figure 4.1 CDSS and the diagnosis process.**

The role of the CDSS program is to help members of the public recognize when they have a problem, in order that they can go to a doctor or to a clinic timeously and be assessed fully and be helped. Figure 4.1 shows a model of the diagnosis process. The diagram has been taken from Chapter 2 Section 2.2.2. The CDSS does not diagnose in the strict medical sense – this is the role of a medical doctor – but it can help the patient identify a problem. In Figure 4.1, the role of the CDSS is described by the first process box. The steps of the other process boxes are done by a doctor. The CDSS program uses symptoms felt by a patient in order to assess whether a patient may have a serious problem. Its action can perhaps be better described as a preliminary diagnosis, which the patient can use to find further help.

A process similar to that employed by a doctor to assess a patient is employed by the CDSS program. Though, its approach uses symptoms only. This is similar to how a consumer health software program such as *Housecall* was derived from an existing software program named *Illiad*, which was meant to assist doctors (Bouhaddou *et al.*, 1998). The doctor may have used *Illiad* when diagnosing a difficult case. For the consumer, *Housecall* or the CDSS can be considered to be relatively sophisticated medical references.

The CDSS program operates by prompting the user (consumer) to enter his/her symptoms. It assesses whether it is likely that the user may be suffering from one of a list of common ailments, such as tuberculosis or hypertension.

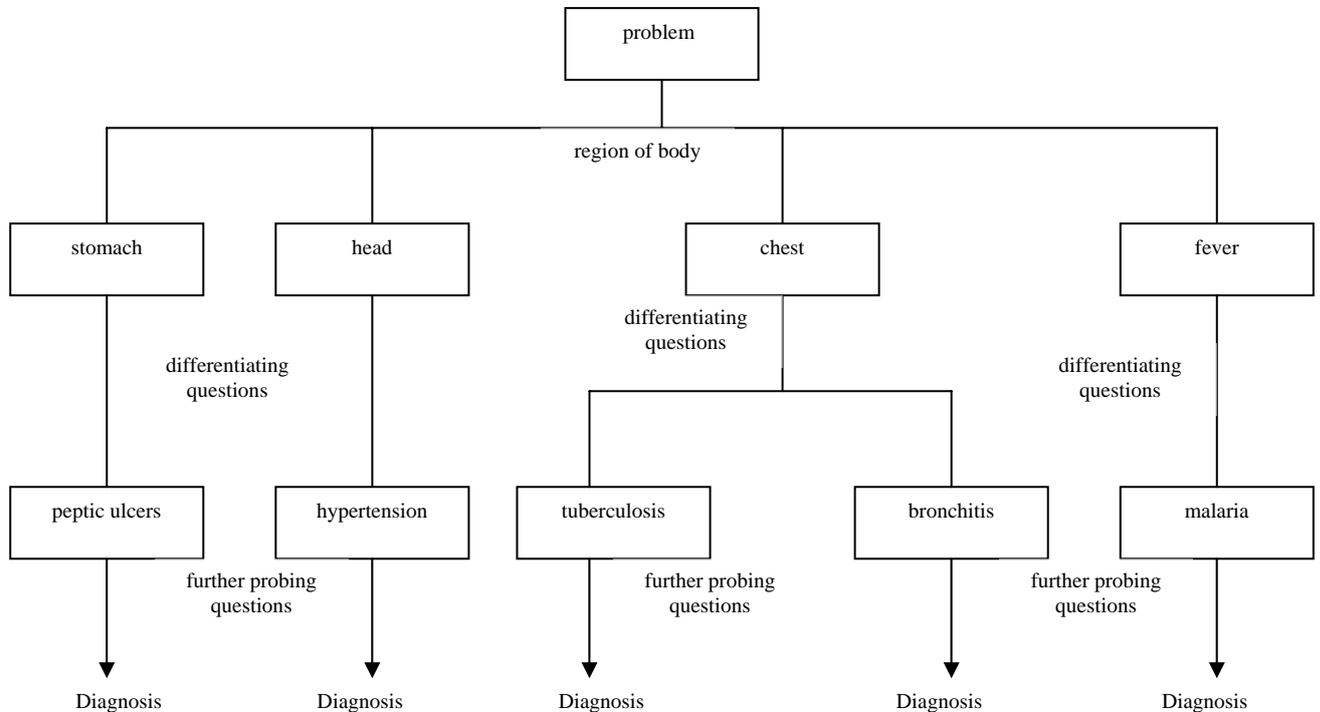
## 4.2 Software Implementation of CDSS

### 4.2.1 CDSS Diagnosis Strategy

The CDSS diagnosis strategy is given by the following steps, which are taken from Chapter 2 Section 2.3.3. Figure 4.2 is a summary of the strategy also taken from Chapter 2.

- Step1. The CDSS enquires about the area of the body in which the problems presents. For example, chest is the area for tuberculosis and bronchitis, while stomach is the area for peptic ulcers. If the patient fails to identify the part of the body in which the problem presents, then the CDSS suggests a fever by asking, in which case malaria is the possibility.
- Step 2. The CDSS checks for certain key symptoms that differentiate illnesses in the part of the body in which they occur, in order to make a tentative diagnosis. For example

tuberculosis and bronchitis are differentiated by weight loss and cigarette smoking. Weight loss would apply more to tuberculosis, while cigarette smoking applies more to bronchitis. Both illnesses are characterized by a persistent cough. If a patient has a persistent cough and has lost weight then tuberculosis is suspected tentatively.

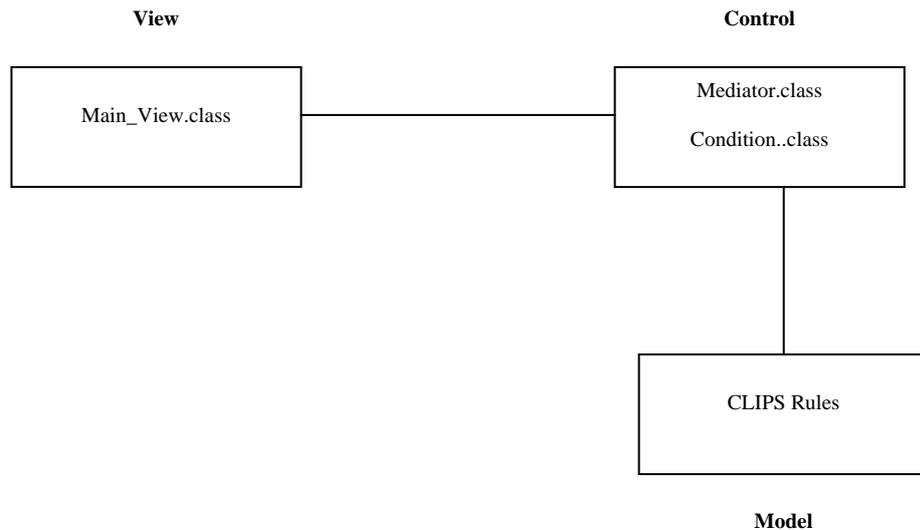


**Figure 4.2 CDSS diagnosis strategy.**

- Step 3. It asks further probing questions which are in line with the standard symptoms for the illness tentatively suggested. This is in order to confirm presence of the illness. For example for tuberculosis it will ask if the person has an appetite, if they experience extreme night sweat, or if they have been in contact with someone with tuberculosis.
- Step 4. At the end of the questions it uses the combination of rules given by the experts to make a diagnosis. For example for tuberculosis the following is a combination {PC WL NS}. It states that if a patient has persistent cough (PC), weight loss (WL) and extreme night sweat (NS) then they may be suspected to have tuberculosis.

## 4.2.2 Software Modules of CDSS

In order to describe the software subparts of the CDSS the MVC (Model View Control) software architecture model has been used. It is normal to split a software application into its main functional parts (Garlan et al, 1994). There are alternative architectural views other than the MVC scheme, but for the purposes of this Chapter the MVC scheme is adequate. Figure 4.3 is a diagram of the CDSS software architecture according to the MVC model.



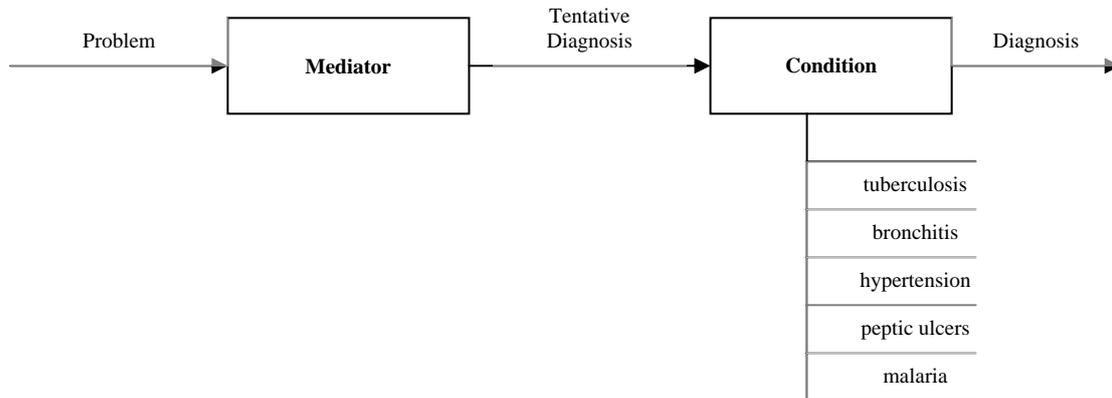
**Figure 4.3 CDSS and the Model View Control scheme.**

The *model* is a knowledge store together with the processes for manipulating it. The *control* operates on the model, like a brain or processor, in order to solve the problem and coordinate the other parts. The *view* is responsible for the dialogue and interface with the user. It gets its feedback from the control.

In the CDSS the model consists of the experts' rules which have been encoded using CLIPS. CLIPS is a production rule language. The model is also called the knowledge base. Control is handled by two Java classes. The first is called *Mediator*, which handles the first two steps of the CDSS diagnosis strategy. Its task is to generate a tentative diagnosis. It then activates *Condition* which is the second module in the control, which handles step 3 and step 4 of the CDSS diagnosis strategy.

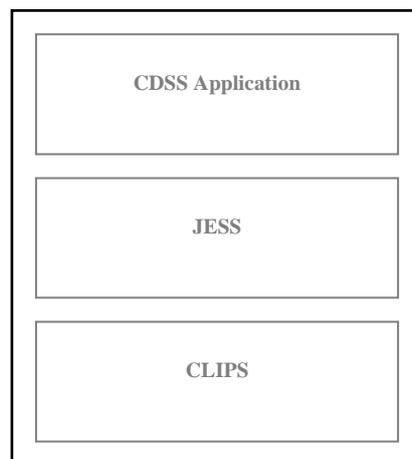
*Condition* has a subclass for each illness of the study. For example if tuberculosis is the tentative diagnosis made at step 2 then it activates the tuberculosis subclass to ask further probing questions in order to confirm presence of illness. At the end of the questions it uses the experts'

rules to make a diagnosis based on the patient’s symptoms. Figure 4.4 shows the diagnosis modules of the CDSS and their interaction. *Mediator* and *Condition* both interact with the knowledge base (model). Another Java class called *Main-view* is responsible for the interface with the user, figure 4.3. It draws its input from *Mediator* and *Condition*. There are other supporting classes in the program. However, the ones mentioned here are the main ones.



**Figure 4.4 CDSS diagnosis modules and their interaction.**

The CDSS program uses CLIPS to encode the experts’ rules. It accesses CLIPS through JESS a Java API which allows a Java program to access the CLIPS rule environment directly. JESS can be classified as a wrapper façade to CLIPS. JESS can be found at: <http://herzberg.ca.sandia.gov/jess/>. Figure 4.5 describes the relationship of the CDSS, CLIPS and JESS.



**Figure 4.5 Relationship of CLIPS, JESS and CDSS.**

## 4.3 User Interface Design

The interface of a program is the point of contact between the user (collaborator) and the program. Though in the current study the interface of the CDSS program is not evaluated, some important guidelines were established from the literature.

### 4.3.1 Criteria of CDSS User Interface

When designing user interfaces, it is often good to use a familiar style so as to take advantage of human recognition. This guideline was followed in the design of the user interface for the CDSS program. In fact, the inspiration for the CDSS interface design was derived from an informal conversation that occurred early on in the project. In that conversation, the CDSS program was referred to as a sort of ATM – a health ATM. An ATM refers to the familiar *automatic teller machine* (auto cash machine). This was found to be a useful idiom to describe the CDSS program interface.

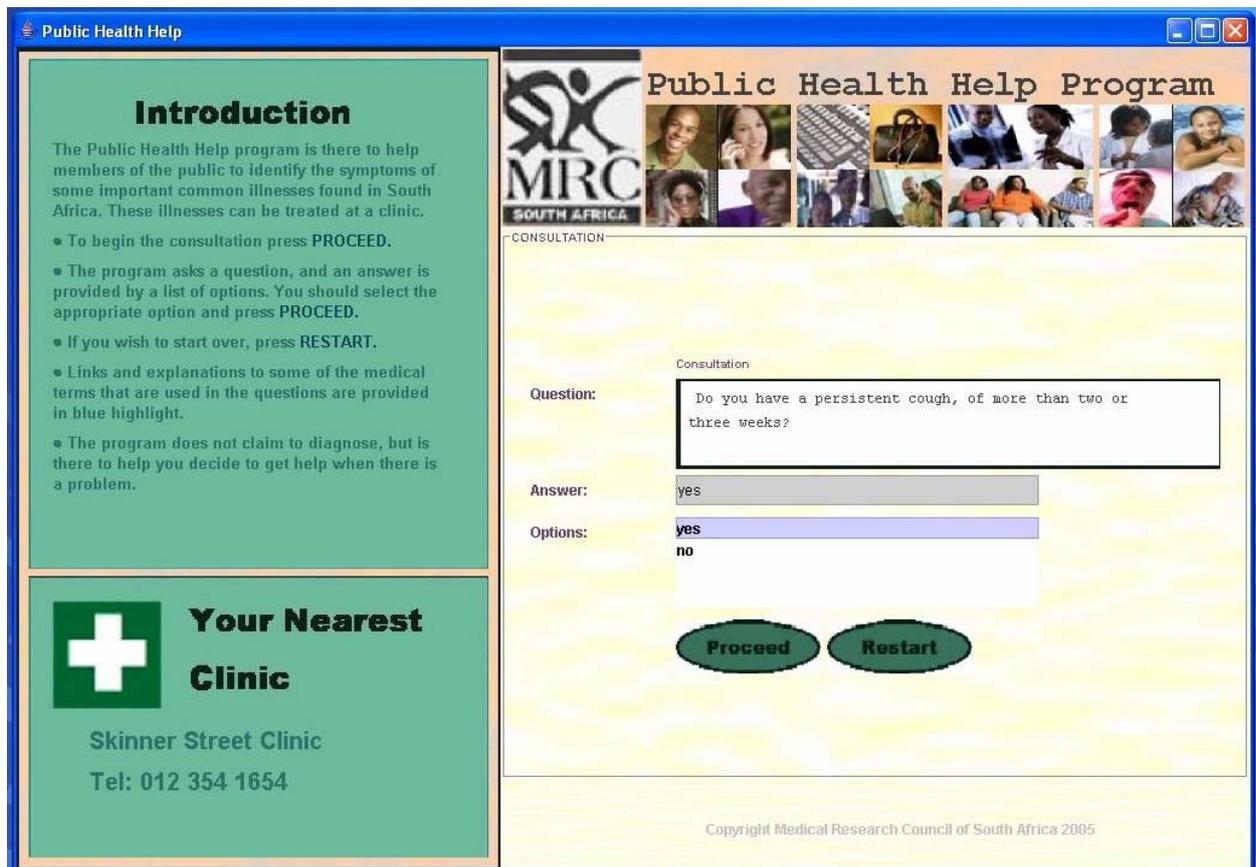


Figure 4.6 Screenshot of CDSS user interface.

Some qualities of an ATM which are important to the CDSS program interface are the following: overall simplicity, few controls and the provision of options. One instance of overall simplicity, is that the operating instructions should be simple. To operate, the user should only have to choose from a few options then receive an answer. Figure 4.6 is a screenshot of the field prototype of the CDSS program which has been developed and which is planned for further evaluation and improvement. For the public it has been called the Public Health Help program, which is a more consumer friendly term than CDSS.

### 4.3.2 User Interface Design Challenges

Capturing symptoms is often the task of a doctor or a health professional. Reasoning about diagnosis often involves the use of medical terminology. A challenge has been how to design the user interface so that information about these two features are made simple and thus become available to a lay member of the public. In trying to meet this requirement the following guidelines were followed:

- An interface was designed based on probing questions used for assessment at primary health care level in South Africa. This questionnaire was created by a nursing sister who was one of the experts in the development team of the CDSS. She used her own experience, literature as well as fresh observation of patient doctor consultation.
- The answers to the questions should either be of the yes / no variety, or should entail simple options that can be provided on the screen.
- Only those symptoms that the patients could identify and describe in everyday language were captured and used. This is something that was decided on in consultation with the experts that were involved in the development of the CDSS program.

## 4.4 Project Team

The development team for the CDSS program consisted of the following specialists.

- A chest specialist who is a member of faculty at Medunsa Medical campus of the University of Limpopo in South Africa.
- A gastroenterologist who is also a member of faculty at Medunsa Medical campus of the University of Limpopo in South Africa.
- A nursing sister with over thirty years experience, including in primary health care.
- And the author, who played the role of knowledge engineer.

The nursing sister created the illnesses questionnaire, with verification and recommendations from the members of medical faculty on the project team. The members of faculty supplied the experts' rules used for diagnosis by the CDSS.

## 4.5 Project Evolution

It is suggested that the prototype of a software program may undergo the following stages:

- Research prototype.
- Field prototype.
- Production model.

These stages were useful for describing the requirements that need to be met when building a product that can be accepted and used in the community. It is envisioned that the CDSS will be developed to be of production strength. By “production strength” is meant that the system will perform effectively at its task and will be safe. This can only be achieved through rigorous and systematic, evaluation and testing. The current study is part of this process. With respect to the above stages, the CDSS prototype that was developed can be classified as being at the stage of transition from a *research prototype* to a *field prototype*.

A research prototype is limited to use by the research team, and if users are involved it is within a very tightly controlled environment. A field prototype is intended for users in the actual deployment environment. It is more sophisticated and safer than a research prototype. The final stage, the production model, has been fully developed and tested, and found to be effective and safe. The relevant support and infrastructure has also been put in place for a successful deployment

The CDSS software program is a comparatively small endeavor which did not demand substantial resources to build. It was built by the knowledge engineer, as a student related project, with input from the other team members at intervals. Perhaps it could also be an example of how to build a software program that is relevant to South Africa. South Africa is a developing country, looking for viable solutions to some of her most pressing problems. The solutions should be developed using limited resources while making a difference.

## 4.6 Conclusion

In conclusion an attempt is made to provide an answer to the question of why the CDSS program is considered an expert system. Two salient features that distinguish it have been identified. Firstly, it has been built using the methodology of knowledge engineering, which involves the use of experts. Secondly rules which are traditionally an expert systems technology were used. They were encoded with CLIPS which is an expert system development language. In creating the CDSS evaluation plan it was also found that strategies used for evaluation of expert systems were applicable. Evaluation of health expert systems is discussed next in Chapter 5.

## Chapter 5

### Health Expert System Evaluation

This chapter covers the theme of evaluation and improvement of rules of the knowledge base of a health expert system. It begins with an introduction of the evaluation of health expert systems which describes the underlying principles to health expert system evaluation. Two case studies are presented. Firstly is the evaluation of MYCIN which shows, by example, the factors of health expert system evaluation. MYCIN is a classic medical expert system. The second case study is the Predictive Toxicology Challenge (PTC), which is an example of the use of the JSM method. The Predictive Toxicology Challenge is part of the inspiration to use methods of FCA in the CDSS study. The chapter ends with a description of the evaluation approach of the CDSS study. The chapter uses the following sub-headings.

- 5.1 Introduction to health expert system evaluation
- 5.2 MYCIN evaluation approach
- 5.3 Predictive Toxicology Challenge
- 5.4 Study designs of knowledge evaluation studies.
- 5.5 Evaluation approach of the CDSS study
- 5.6 Conclusion

### 5.1 Introduction to Health Expert System Evaluation

#### 5.1.1 Kinds of Evaluation

There are two important general evaluation dimensions of an expert system. The first point of evaluation checks if the program is reliable and accurate (Berners, 1999b:61). This entails checking that the knowledge base is accurate and that the advice of the program is reliable. The second important point of evaluation is checking the social impact of the program. For example, the social impact study for the CDSS program would judge whether or not the CDSS program makes a difference in the health of the community. The current study was meant to evaluate the knowledge base and is part of the first dimension of evaluation of expert systems. The social impact is beyond the scope of the present study, and will be the subject of a further study.

When building an expert system there is a degree of informal testing that happens. This occurs when the developer enters new rules into the program. The developer would normally run a set of test cases to check that the system still functions correctly, and that the new rules work properly. The following is a list of some of the evaluations that are often done for expert systems (Buchanan & Shortliffe, 1984c:571).

- System Accuracy and Advice
- Correct Reasoning
- The Human Computer Interface
- Hardware and Efficiency Analysis
- Social Impact

It is clearly the case that evaluation is not an end in itself. In each of the above evaluation processes, the purpose is to identify weaknesses and to ensure that corrective action is taken. This will often result in an incremental process, whereby changes are made, the revised system evaluated, further weaknesses identified, improvements made, etc.

#### 5.1.1.1 System Accuracy and Advice

*System accuracy and advice* needs to be evaluated. This entails an evaluation of both the knowledge base and of the performance of the system at giving advice. The aim this study was to explore the utility of FCA methods to evaluate the rules of a knowledge base of a health expert system. Consequently the performance was also measured. So it was possible to make an assertion about the reliability of the advice of the CDSS, and hence its potential usefulness.

#### 5.1.1.2 Correct Reasoning

*Correct reasoning* considers the process used by the expert system to arrive at conclusions. The correct reasoning was discussed under the diagnosis strategy in Chapter 2 in Section 2.3.3. In the discussion it was shown how the knowledge base was used in the diagnosis strategy of the CDSS.

#### 5.1.1.3 The Human Computer Interface

The *human computer interface* needs to be evaluated. This is the subject of a planned study. One of the human computer interface evaluations to do is to check the appropriateness of the medical language used for members of the public. In South Africa, the target users are not first language English speakers, so there are issues of translation. There is also low computer literacy

among the target users. These issues need to be resolved in order for the full benefits to be realised.

#### 5.1.1.4 Hardware and Efficiency Analysis

*Hardware and efficiency analysis* may also be done. If, for example, the program uses network resources, it may be important to evaluate the extent to which these resources are used. The envisaged complete CDSS system is not intended to be on a network, and it is unlikely to consume significant other computer resources. Rather, it is seen as a stand-alone system, running on a fairly standard PC. Of course, in the much longer term, it may evolve into a networked system where the collection of statistical information becomes a priority, but such matters are beyond the scope of the current study.

#### 5.1.1.5 Social Impact

The *social impact* of the CDSS is not part of the current study but is planned for a later phase. Again, in the longer term, it is clearly hoped that the system will become a real asset to health consumers, and add value to the health resources of the country. In Chapter 2 the motivation for developing the CDSS was given though.

### 5.1.2 Factors in Health Expert System Evaluation

This section explains the issues to consider in evaluation of expert systems (Buchanan & Shortliffe, 1984c:579).

In order to test the performance of the program an *objective standard* needs to be devised against which the advice of the program will be compared. This is sometimes called a *gold standard*. When a new medical diagnostic procedure is introduced the *gold standard* is often a laboratory test, or an invasive procedure like a biopsy, or the results of an autopsy. There are two main options for a gold standard in evaluation of expert system. The first option is to use what eventually turns out to be confirmed as correct either by carrying out an invasive procedure or by some other method. The second option is to use what a human expert concludes on a case. Sometimes this is the only option available. In the study the diagnosis of the treating doctor was used as the standard for measuring performance.

Sources of *bias* need to be considered. This may necessitate *blinding* measures. In the next section on the discussion on the MYCIN evaluation test case, an example of how bias is handled is provided.

The evaluation often should happen in stages. Each stage should evaluate a part of the program. For example, evaluating system accuracy and reliability is often the first task. Then the system can be tested with users. If, for example, the system performs optimally at decision making, but users experience problems with it, then it can be assumed that the reason for these problems is something else other than the decision-making ability of the program. The CDSS, evaluation began with the evaluation of the knowledge base which is the subject of the study.

The other evaluation factor that should be considered is the definition of a *realistic standard of performance* for the program. For example, the CDSS program may not catch all instances of people that have the target illnesses, but the numbers that are captured should be significant. The question that needs to be asked is: What should this significant performance value be? For example, it was considered that a 90% performance rate would be too high, considering the fact that the CDSS program uses patient symptoms only in its assessment. Consequently it was decided to aim for a performance rate of approximately 70%. This was inspired by the study on Housecall (Bouhaddou et al, 1998). Housecall is a commercial consumer decision support system for illnesses of internal medicine. The conclusions of Housecall were restricted to 70% certainty.

## **5.2 Evaluation Approach of MYCIN**

Evaluation of MYCIN (Yu *et al*, 1984:589) is an example of how evaluation of a health expert system could be done. Its focus is on evaluating the advice of MYCIN rather than the knowledge base. But it shows practically some of the factors that are considered in health expert system evaluation.

### **5.2.1 Domain of MYCIN**

MYCIN's domain of operation was the treatment of blood infections. The aim of MYCIN was to help physicians make the most appropriate treatment for an agent which had infected the blood of a patient. Some drugs are effective against certain agents, while against others they are not effective. Some which are effective are also very toxic. For example, they may only be used

in laboratory experiments where it is safe or in chemotherapy where desperate measures need to be taken. Therefore, the role of choosing a treatment for blood infection is not trivial and a system of this kind could be helpful.

### **5.2.2 MYCIN Evaluation Plan**

The evaluation of MYCIN's advice was done as follows: Ten patients with infectious meningitis were selected by a physician that was not acquainted with MYCIN. The cases were considered difficult cases and were selected to cover a wide range of meningitis illnesses. For example, some were fungal, some were viral and others were bacterial. A detailed clinical summary was made of each case.

Eight experts in medicine and infectious disease were selected. They included five senior faculty members at Stanford University Division of Infectious Disease, one senior resident in medicine, a senior medical student, and a senior post doctoral student in the Department of Infectious Disease. Each expert was given the clinical summaries and asked to make a prescription for each. The clinical summary data was also used on MYCIN, and MYCIN also wrote its own prescription, on the ten cases. The prescriptions given by the treating physician were also recorded. For each patient there were ten prescriptions: eight from the panel of experts, one from MYCIN, and one from the physician actually treating the patient. All those that wrote a prescription (including MYCIN) in the study were called prescribers.

The prescriptions were given to another panel of experts who came from outside of Stanford University. These experts were called the evaluators. This panel consisted of eight experts in infectious diseases. Their task was to evaluate the prescriptions that had been made. The evaluators were not told which prescription came from MYCIN, which came from the faculty, student or treating physician. They were asked to evaluate each prescription, then to assign a score out of 80. In this evaluation the gold standard was the panel of expert evaluators. To summarise: there were eight evaluators, who had to evaluate ten prescribers who had to assess and write prescriptions for ten meningitis patients. Table 5.1 summarises the results.

### **5.2.3 MYCIN Evaluation Results**

According to the results, the performance of MYCIN compared favourably to that of the infectious disease experts of Stanford University.

**Table 5.1 Results of MYCIN evaluation.**

<u>Prescribers</u>	<u>Score</u> <i>(total - 80)</i>	<u>Failed Therapies</u>
MYCIN	52	0
Faculty-1	50	1
Faculty-2	48	1
Infec. dis. fellow	48	1
Faculty-3	46	0
Actual Therapy	46	0
Faculty-4	44	0
Resident	36	1
Faculty-5	34	0
Med. Student	24	3

MYCIN actually had the highest score, and did not provide a failed therapy. This limited study was able to give an indication of the potential of MYCIN. It could perhaps have been used as support for a wider more comprehensive study which would have a larger sample of patient cases but which would be harder and more costly to do. An initial small sample study was also part of the evaluation strategy of the CDSS project. Next is the Predictive Toxicology Challenge which is another example of an evaluation study.

### **5.3 Evaluation Approach of Toxicology Analysis by Means of the JSM Method**

The following case study is an example of use of the FCA formulation of the JSM method. This study will be referred to as the Predictive Toxicology Challenge (PTC). The PTC was part of the inspiration to use FCA in the study.

#### **5.3.1 Domain of Predictive Toxicology Challenge**

The PTC used the JSM method to predict the toxicity of chemical substances. In the PTC, data was supplied of known toxic and non-toxic substances, and their chemical structure

combinations. Based on this data the JSM method was used to identify chemical substructures which were predictors of toxicity, which can be used in general to decide toxicity of substances. This problem is important in that some substances are needed for daily use either for industrial or domestic purposes. Their toxicity needs to be evaluated as some toxins may cause certain cancers or other problems. Toxicity is often determined by doing trials which can take a very long time and be costly. Hence a reliable method that predicts toxicity based on chemical structure can be beneficial.

The character of weak bonds of a compound can be used to determine toxicity. A lot of knowledge about chemistry is not needed in order to understand the basic idea behind this experiment. Examples of weak bonds are double bonds in a carbon chain (for those who are familiar with organic chemistry). There are many other examples that can be found in chemistry. The nature of weak bonds depends on  $\pi$ -electrons of the compound. There is a language (scheme) which is called FCSS (*fragmentary code of substructure superposition*) which is used in the classification of these weak bonds. It was first proposed by Avidon and Pomerantsev (1982). By means of this language a chemical compound can be represented as a set of substructures that are centers of  $\pi$ -electrons. In this discussion these chemical substructures have been called FCSS combinations, though this may not be the strict term used in the FCSS domain. In the PTC study they were also referred to as features of a chemical compound (Predictive Toxicology Challenge, 2001). A chemical compound can be described by an FCSS combination.

### 5.3.2 PTC Evaluation Plan

A dataset of substances whose toxicity was known was provided to the PTC. There were about 780 substances in the dataset. Certain of these substances were known to be toxic, according to the JSM method these were called positive examples. While others were known to be non-toxic, these were called negative examples. The toxicity of these substances had been tested on rats and mice, of male and female. For example, some substances are toxic against rats but not mice. The gender is also a factor, whether female or male. There are various combinations of toxic agents against rats and mice, of male and female. The FCSS structure of each substance in the PTC dataset was known. This dataset was split into two. There was a training dataset and test dataset. The training dataset was used by the JSM method to identify FCSS combinations which could be used as good predictors of toxicity and non-toxicity. Then the set of FCSS combinations

identified by the JSM method were tested on the test dataset to find out if they were able to make good predictions or not.

The JSM method generates hypotheses based on positive and negative examples which have been supplied. In the case of the PTC, hypotheses are FCSS combinations which predict toxicity and non-toxicity. Hypotheses which are used to make prediction of toxicity are called positive hypotheses. These are hypotheses which are only found among positive examples. While negative hypotheses predict non-toxicity and are only found among negative examples. Chapter 7 describes in detail the JSM method. In this chapter the reader should understand the role of the JSM method within the evaluation strategy.

This was applicable to the CDSS study because the patient data that was collected can be regarded as examples. Those patients who are diagnosed with tuberculosis are considered positive examples of tuberculosis. Conversely, those patients who have some symptoms of tuberculosis such as coughing, but do not have tuberculosis are negative examples. Such a person may have acute bronchitis, for example. Based on positive and negative examples the JSM method can be used to generate hypotheses which can then be used to predict if someone has (is suspected to have) tuberculosis based on symptoms. The hypotheses are in fact combinations of patient symptoms which may predict illness.

### **5.3.3 Results of PTC**

The details of the results of the PTC can be found in Blinova *et al* (2002). It is of value in that it showed that the JSM method had good performance, according to the authors. They state that it was rather conservative. This discussion is continued in the conclusion to Chapter 8. Here it can be said that conservative suggests that more evidence may be required before a prediction, hence fewer predictions result. But all predictions made tend to be correct.

## **5.4 Study Designs of Knowledge Base Evaluation**

There are a number of study designs that can be followed in the evaluation of the knowledge base of an expert system. These study design approaches also use a various number of data analysis techniques. An example of a knowledge base evaluation study design which was found in the literature is the following (Verification, Validation and Evaluation, n.d.). The knowledge base was evaluated by a team of experts as follows. Knowledge items of the knowledge base

were given to a team of experts to evaluate. This team shall be called the expert evaluators. A knowledge item could be a production rule or any important assumption of the knowledge base. The expert evaluators consisted of eight experts. They were asked to evaluate whether each knowledge item was true or not. If a majority of expert evaluators agreed on an item then it was assumed to be true, and it was given a confidence rating based on their ratings. A statistical analysis approach was used to analyse the results. The example of the evaluation approach of MYCIN was also dependent on expert evaluators to score the diagnoses that were made. The CDSS evaluation approach follows next. In the CDSS study the expertise used in the knowledge base is evaluated against a sample of actual field data using FCA.

## **5.5 Evaluation Approach of CDSS Study**

The aim of the CDSS study was to test the rules of the knowledge base of the CDSS against an actual sample of patient data using FCA methods. The results of this could be used to decide the utility of FCA at evaluating and improvement of the knowledge base of rules of an expert system. The study design was as follows.

The field prototype of the CDSS program was deployed in a primary health care (PHC) clinic where it was used to collect data on patients' symptoms. A PHC clinic was chosen as the data collection site as it would be easy to find patients with the symptoms and illnesses of the study. Study patients were chosen from among patients who had come for a visit to the clinic. Patient data was collected as follows. The CDSS was used to collect the patient's symptoms using the illnesses questionnaire. It was possible to use the CDSS to collect patient symptoms because it used the same questionnaire that would be used by the researcher. Thereafter the patient went to see the doctor. Then the doctor completed the patient's data by providing the diagnosis. The data for each illness was formatted into a table which contained the symptoms for each patient of the study as well as the treating doctor's diagnosis. Then the FCA methods were used to generate rules from the patients' data. The rules generated by the FCA methods were then compared to rules provided by the experts. The new rules generated by FCA method were also tested on the CDSS to check whether their inclusion improved the performance of the CDSS or not.

The PTC study was the inspiration for exploring the utility of methods of FCA in the CDSS project. In addition to JSM hypotheses it was found that FCA implications could also be used. Then it was decided to use both methods. It was thought that the two methods may complement

each other, and that they could be compared. FCA implications and JSM method hypotheses are equivalent to rules. Hence use of the term that FCA implications and the JSM method are used to *generate rules*. The rules generated by the FCA methods were compared to those that were given by the experts as follows.

- *Rules Subsumed*. How many experts' rules have been subsumed by the FCA rules? These can be considered to be experts' rules which are accounted for by the FCA generated rules.
- *New Rules*. These are rules generated by the FCA methods which have no corresponding experts' rules. These rules may be considered enrichment to the CDSS if the experts accept them as valid.
- *Rules not Accounted*. These are expert rules which have no corresponding FCA rules as defined above in the first bullet. These rules can be considered misses (omission) by the FCA methods.

The new rules generated by the FCA methods were also tested by measuring the performance of the CDSS at giving advice. By performance the following is meant. How many times did the diagnosis of the treating doctor coincide with that of the CDSS. For example, if out of ten cases it makes the same diagnosis as the treating doctor for seven cases, then the performance of the CDSS will be given as 70%. Its performance was measured twice. First it was measured with the original experts' rules only. Then the new rules generated by FCA were, and the performance was measured again. The aim was to check whether performance improved or not, as a result of inclusion of the new FCA generated rules. To do this the data was split into two sets. There was a *training set* which was used for generating the FCA rules. And there was the *test dataset* which was used to test its performance with the new rules included, against the original experts' rules. The two datasets were disjoint, this is because the new rules would have a superficially high performance against the training dataset from which they were derived. Splitting of the data was random. The training dataset was collected first, then the test dataset. How this was done is explained next in Chapter 6 on data collection. The results of data analysis were then taken back to the experts for consideration in the iterative manner of knowledge engineering of expert systems.

Consequently its performance was used to give an indication of the reliability of its advice, and hence its potential usefulness.

## 5.6 Conclusion

It is important to evaluate an expert system, to make sure that it does its task satisfactorily and that it will be useful. In the study the knowledge base of the CDSS was evaluated using FCA methods. This was done in order to identify weaknesses and to improve the quality of the knowledge base. Chapter 6 next, describes the data collection process. Chapter 7 describes the FCA data analysis methods in detail, and Chapter 8 provides the actual results.

## Chapter 6

### Data Collection

This chapter describes the process of data collection. The source of the data is explained, the choice of study site, the sampling strategy, and the formatting of the data. The chapter has the following subheadings.

- 6.1 Source of Data
- 6.2 Patient Sampling
- 6.3 Study Patient Data Format
- 6.4 Conclusion

#### 6.1 Source of Data

In the beginning there were two main options for acquiring the data that was needed. The first was to find an existing dataset of existing patient records. From these patient records, symptoms and diagnosis by the doctor would be extracted. Then FCA methods would be used to generate rules from the data. The alternative was that the data could be self-collected from patients. The first option was judged to be problematic, not only because of the logistical difficulties of gaining access to a set of pre-existing records, but also because the existing paper based records of the public health system which document patient symptoms and diagnoses, do so differently to the way required by the study. Although an electronic health record database may have been better, such a facility does not yet exist in the public health system in South Africa. The second option was therefore chosen.

A PHC clinic was chosen as the data collection site as it would be easy to find patients with the symptoms and illnesses of the study. A district PHC clinic in Pretoria city centre was chosen, because it covered all the illnesses of the project, except for peptic ulcers. There are also medical doctors in attendance at the Skinner Street Clinic, which may not be the case for most PHC clinics. For peptic ulcers, it was intended that the gastroenterological unit at Pretoria Academic Hospital be used, where peptic ulcer patients could more easily be found. However peptic ulcers data was not collected because the gastro-enterological unit of Pretoria Academic Hospital was in the process of relocating during the period of data collection. Malaria data was

also not collected because malaria cases were not found in the study clinic. This is because of the season chosen for the study and because in Gauteng, Malaria is not very common. Though malaria and peptic ulcers data was not included, the results of data analysis which are discussed in Chapter 8 were found to be adequate to decide the utility of FCA at evaluation of the CDSS.

## 6.2 Patient Sampling

The initial aim was to identify 20 patients per study illness. For each illness this data would be split into two sets. Firstly the training set which was used for generating rules with the FCA methods. Secondly the test dataset which was used to test the rules that had been generated with the FCA methods. Each set would consist of positive examples and negative examples. An explanation follows of the terms negative and positive examples. Those patients who have tuberculosis like symptoms and are diagnosed with tuberculosis are considered positive examples. Those that have tuberculosis like symptoms but who are not diagnosed with tuberculosis are considered negative examples. For example they may have acute bronchitis or some other chest infection that may cause persistent cough. It was decided that both datasets should consist of at least 30% of negative examples each. The datasets were decided randomly in the following way. The data for the study was collected in two cycles. The first cycle collected the training dataset, and the second cycle collected the test dataset. According to time and resources available the sampling strategy suggested in this paragraph was considered reasonable.

The size of the sample was also inspired by a study done by Eminovic et al (2004) which was a preliminary study on a patient triage system which was done over an Internet chat facility. In that study the patients communicated with a PHC nurse using an Internet chat, and afterwards the nurse made a recommendation, even though he/she had not physically examined the patients. The data collection was done over two weeks, and it had about 40 patients as subjects. Preliminary findings were made from its results about the triage system. Next it is explained how the positive and negative examples were gathered.

As explained, data was collected only for three conditions: tuberculosis, bronchitis and hypertension. It was initially decided to select patients according to their main complaint so as to maximize the chances of getting a sufficient number of positive examples. The number of positive examples per set was set at about 70%. Since it was only possible to establish *ex post facto* whether or not an initially interviewed patient actually had one of the conditions, it was not possible to guarantee that interviewed patients would turn out to be positive exemplars of these

conditions. If patients were to be selected randomly the researcher who was working alone would have to spend a very long time at the clinic in order to find the required number of subject patients. Then the following was decided. It had been observed that there were many hypertension patients who were already diagnosed, who were only visiting the clinic to collect medication and for scheduled visits. These already diagnosed chronic cases were easy to identify than the random cases present in the queue which still needed to be diagnosed. It was decided to select from these known chronic hypertensive cases. The interview was then conducted by asking the patients how they felt when they first sought help. Rather than how they felt at the time of the study interview. This is because their conditions may have improved by the time of the study as a result of medication. So in this approach it was relied on the ability of the subject patients to remember how they felt when initially ill. The same was done for tuberculosis and bronchitis. Subjects were chosen from among those who were already diagnosed, and who were coming for scheduled visits as well as medication. This was done with the help of clinic staff.

The negative examples were found by identifying patients who had some of the symptoms of the illness for which data was being collected, but who were not diagnosed with the illness after assessment by the doctor. For example to find negative examples for tuberculosis, patients who had persistent cough and perhaps weight loss but who were not diagnosed with tuberculosis by the treating doctor were sought. For bronchitis patients who had persistent cough and who smoked cigarettes but were not diagnosed with chronic bronchitis were sought. For hypertension, patients with persistent headache or vision problems but who were not diagnosed with hypertension were sought. These patients were sought with reliance from the treating doctors who would give the researcher the file of such patients, so that the researcher could interview them.

The study was also limited to patients who were literate enough in English in order to answer the CDSS questionnaire. The researcher entered the patients' symptoms into the CDSS, rather than the patients themselves to use the CDSS program. This was because it was found that the levels of computer literacy were low among patients of the target clinic.

A patient went through the following steps when taking part in the study.

- Patient consent was obtained before participation in the study.
- Patient was assigned a study number, and given a *study form* with the patient's study number. The aim of the study form was to identify the patient to the doctor as a participant in the study.
- The patient history was collected by the researcher via the CDSS questionnaire.

- The patient next saw the doctor and handed over the study form.
- After the consultation, the doctor placed the study patient's file aside for the researcher to use.

The researcher accessed the study patient's record in order to record the conclusions reached by the doctor. The patients' records were not removed from the clinic. Table 6.1 shows the number of patients that were collected.

**Table 6.1 Number of study patients collected.**

Tuberculosis	31 patients
Chronic Bronchitis	7 patients
Hypertension	21 patients

For bronchitis twenty (20) study patients were not found, though initially that had been planned. This is because diagnosed cases of chronic bronchitis were hard to find in the study clinic. Perhaps as had been planned for peptic ulcers, Pretoria Academic hospital should have been used to collect bronchitis data. However the limited sample of bronchitis patients that was ultimately collected was used. It was found that it did add meaningfully to the outcomes of the study. There were more subjects patients collected for tuberculosis than was planned, and this was accepted. The ratio of 30% negative examples in the data was not always achieved. But as this was an initial limited study it was decided to proceed, and it was found that the results of data analysis did satisfy the aims of the study.

A study protocol was submitted to the ethics committee of the University of Pretoria Health Sciences faculty. The study was approved. Permission to use the clinic was also required. In this regard an application to the provincial Gauteng Health Department was made. Permission for this was also granted. In addition to project ethics review and application for permission from the province and district health department offices, administration also had to be done at the clinic. This entailed meeting the clinic manager for go ahead. Introduction and

**Table 6.2 Data collection schedule.**

<u>Activity</u>	<u>Duration</u>	<u>Remark</u>
Ethics Review	4 weeks	<i>Wait for reply</i>
Permission Provincial Health Department	2 weeks	<i>Wait for reply</i>
Permission District Health Department	1 week	<i>Wait for reply</i>
Clinic Administration	1 week	<i>6 hours per day</i>
Data Collection	6 weeks	<i>6 hours per day</i>

meeting of clinic doctors and nursing sisters and other staff that helped in the study. A confidential consulting room also had to be found.

Data collection was done by the author of the dissertation alone. Administratively he was helped by clinic staff. Table 6.2 shows the summary of activities of data collection and their duration.

### 6.3 Formatting of Study Patient Data

The data collected in the study was formatted into tables. An example is shown below for chronic bronchitis.

*Data Field Keys:* PC (persistent cough), SP (sputum production), MC (sputum mucopurulent), CS (clear sputum), CT (chest tightness), SB (shortness of breath), WC (wheezing chest), SM (smoking), CBr (chronic bronchitis).

**Table 6.3 Example, formatting of study patient data.**

	<b>PC</b>	<b>SP</b>	<b>MC</b>	<b>CS</b>	<b>CT</b>	<b>SB</b>	<b>WC</b>	<b>SM</b>	<b>CBr</b>
Patient-1	X	X	X					X	
Patient-2	X							X	
Patient-3	X	X		X	X			X	
Patient-4	X	X	X		X	X	X	X	X

The table can be read as follows, Patient-4 has the following symptoms: persistent cough (PC), produces sputum (SP), the sputum is mucopurulent (MC), there is chest tightness (CT), shortness of breath (SB), a wheezing chest, and smokes cigarettes (SM). A cross marks the

presence of a symptom. The patient symptoms are entered into the CDSS program by the researcher. The diagnosis for Patient-4 is chronic bronchitis (CBr). The diagnosis is provided by the treating doctor. It is shown on the last column, and it is called the target attribute. A cross marks the presence of illness. In this case it is chronic bronchitis. Only Patient-4 has chronic bronchitis, and is considered a positive example. The other patients are negative examples. This table can easily be converted into a CSV (comma separated values) data type which can be input into the automated data analysis tools of the project.

In addition to the treating doctor's diagnoses, the CDSS program made its own conclusions. The CDSS diagnoses were stored internally in the CDSS program and the researcher could access them in order to make the performance comparison of the study. The conclusions of the CDSS though were not shown to the doctor or to the patients.

## 6.4 Conclusion

Data on patients' symptoms and doctor's diagnoses was collected. A PHC clinic was chosen for data collection because that is where people go for the illnesses of the study, but the actual deployment of the CDSS may be away from clinics and doctors' facilities in order to help members of the public decide to seek help. Chapter 7 next discusses the data analysis techniques, and Chapter 8 is on data analysis.

## Chapter 7

### FCA Data Analysis Techniques

*FCA implications* and the FCA formulation of the *JSM method* are the data analysis techniques used in the study. The FCA formulation of the JSM method has been referred to simply as the JSM method. The two methods are part of formal concept analysis (FCA). This chapter explains the theoretical foundations of FCA and of the two FCA methods. The following basic definitions of FCA are given: a *formal context*, *formal concept*, the *prime operation*, and a *formal concept lattice*. FCA implications are discussed first, then the JSM method. Automated tools that were used to generate implications and hypotheses are also presented. The chapter ends with some examples of FCA implications and JSM hypotheses. It has the following subheadings:

- 7.1 Introduction to FCA
- 7.2 Implications
- 7.3 The JSM method
- 7.4 Example of rules using implications and hypotheses
- 7.5 Conclusion

### 7.1 Introduction to Formal Concept Analysis (FCA)

#### 7.1.1 History of FCA

Formal concept analysis (FCA) was developed around 1980 by R. Wille. In the years after that, R. Wille worked with members of the Research Group on Formal Concept Analysis at the Technische Hochschule Darmstadt Germany to develop it further. The main text book on FCA is by R. Wille and B. Ganter (1999). R. Wille was inspired by the work of mathematician G. Birkhoff on lattice theory. The work of Birkhoff dates to 1940. FCA has an annual international conference, ICFCA (International Conference on Formal Concept Analysis). Its application areas include Artificial Intelligence, machine learning, information retrieval and discovery of association rules. In the study it was used to generate rules from patient data, which perhaps can be called rule discovery.

The central notion in FCA is a formal concept, which is called simply a concept. The adjective formal is added to denote that this is mathematical construct rather than the everyday understanding of the term. Before giving the definition of a concept, a formal context and the prime operation are defined. An example of a context is also given.

## 7.1.2 Definition of a Formal Context

**Definition 7.1** (Ganter & Wille, 1999)

A formal context  $K = (G, M, I)$  consists of a set of objects  $G$  and a set of attributes  $M$  shared by these objects.  $I$  is called the incidence relation, it is used to link an object with attributes.  $gIm$  is used to denote that object  $g$  has attribute  $m$ , i.e. attribute  $m$  is incident with  $g$ . The relation could also be represented as a tuple  $(g, m)$  and is equivalent to  $gIm$ . An example is made in Table 7.1.

□ {End of definition 7.1, the end of a definition or a proposition is delimited by the following symbol, □.}

**Table 7.1 Context: “Some Life and Water”.**

	need water	live land	live water
Corn	X	X	
Dog	X	X	
Fish	X		X

A cross table is often used to represent a context. Cross Table 7.1 is of a context called *some life and water*. This is a simple example created by the author. An X is used to represent an incidence relationship between an object and an attribute. The objects are listed in the rows, while attributes are listed in the columns. For example, corn needs water and lives on land (grows on land), while a fish needs water and lives in water.

### 7.1.3 Definition of the Single Prime Operation

In FCA the single prime operation is the basic operation. Its operands are either a set of attributes or a set of objects.

**Definition 7.2** (Ganter & Wille, 1999)

For a set of objects  $A$  of a context  $K = (G, M, I)$ ,  $A \subseteq G$ ,  $A'$  defines the set of attributes common to the objects of  $A$  as follows:

$$A' = \{m \in M \mid gIm \text{ for all } g \in A\}$$

Likewise  $B'$  is defined for  $B \subseteq M$ , as follows: (It is the set of all objects of  $K$  which have the attributes of  $B$ .)

$$B' = \{g \in G \mid gIm \text{ for all } m \in B\}$$

□

For example, in the *some life and water* context. If  $A = \{\text{corn, dog}\}$ , then the set of attributes common to them is  $A' = \{\text{need water, live land}\}$ . Note that  $A' \neq \{\text{live land}\}$ . This is because this is not the only attribute common to  $A$ , need water must also be included. If  $B = \{\text{need water}\}$ , then the objects that have this attribute are  $B' = \{\text{corn, dog, fish}\}$

### 7.1.4 Definition of a Concept

The intuition behind a concept of FCA is derived from the everyday understanding of the term. Philosophers, such as John Stuart Mill in the nineteenth century and Francis Bacon in the sixteenth century, and as far back as Plato, have suggested formal systems of logic whereby a concept has objects associated with it and a set of attributes shared by the objects. For example, in the *some life and water* context (formal) the following is a concept:  $(\{\text{Corn, Dog}\}, \{\text{need water, live land}\})$ . The objects of the concept are *Corn* and *Dog* and the attributes that they share exactly are *need water* and *live on land*.

**Definition 7.3** (Ganter & Wille, 1999)

A formal concept of a (formal) context  $K = (G, M, I)$  is a pair  $(A, B)$  with  $A \subseteq G$  and  $B \subseteq M$ ,  $A' = B$  and  $B' = A$ .  $A$  is called the extent of the concept and  $B$  is called the intent.

□

For the concept  $(\{\text{Corn, Dog}\}, \{\text{need water, live land}\})$ .  $\{\text{Corn, Dog}\}' = \{\text{need water, live land}\}$ , and  $\{\text{need water, live land}\}' = \{\text{Corn, Dog}\}$ . The objects of the concept are called the *extent*, and the attributes the *intent*. The following proposition is based on this operation. It introduces the double prime operator. The double prime operator is simply priming what has been primed already. It is called the *closure* of a set. The definition of a closure system is given in Definition 7.6 below.

**Proposition 7.1** (Ganter & Wille, 1999)

For a context  $K = (G, M, I)$  with  $A \subseteq G$  and  $B \subseteq M$  the following can be said:

$$A \subseteq A'' \text{ and } B \subseteq B''$$

From this proposition the following suggestion can be made: for a context  $K = (G, M, I)$  with  $A \subseteq G$  and  $B \subseteq M$ , the following can be said:  $(A'', A')$  and  $(B', B'')$  are valid concepts of the context  $K$ . Further it can be said that  $A$  is an extent if and only if (iff)  $A = A''$  and  $B$  is an intent iff  $B = B''$ .

□

For example, in the *life and some water* context if  $A = \{\text{Corn}\}$ , then  $A' = \{\text{need water, live land}\}$ .  $A'' = \{A'\}' = \{\text{need water, live land}\}' = \{\text{Corn, Dog}\}$ .

**7.1.5 Definition of a Lattice**

The concepts of a context can be ordered using the following scheme (Ganter & Wille, 1999). If  $(A_1, B_1)$  and  $(A_2, B_2)$  are concepts of a context, then  $(A_2, B_2)$  is a sub-concept of  $(A_1, B_1)$  iff  $A_2 \subseteq A_1$ , this is equivalent to  $B_1 \subseteq B_2$ . This can be interpreted to mean that the more general concept includes more objects, while the more specialised concept has fewer objects. But the

most general object has the least attributes common. In general, the bigger the set then the fewer the things common.

**Definition 7.4** (Ganter & Wille, 1999)

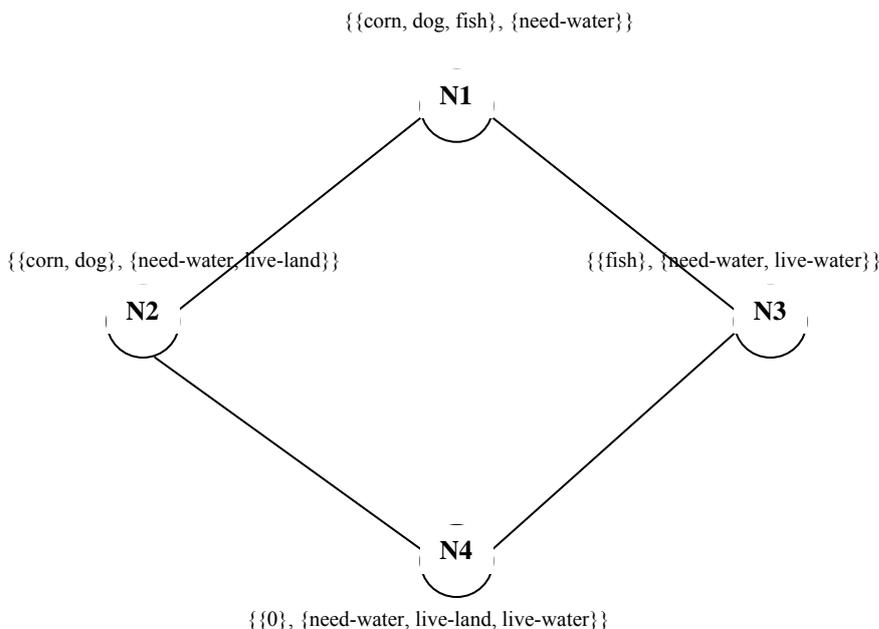
A *formal concept lattice* of a context  $K = (G, M, I)$  is defined as the set of all concepts of a lattice and their ordering relation. It is denoted  $\mathfrak{B}(G, M, I)$ . The *fundamental theorem of FCA* (Ganter & Wille, 1999) states that for any subset of concepts of a concept lattice there is a *supremum* and an *infimum*. The *infimum* is also called the *meet* operation, and *supremum* the *join* operation.

□

**Definition 7.5** (Ganter & Wille, 1999)

The *infimum* of a set of concepts is the greatest concept of the context that is less than or equal to all the concepts in this set. It can be said to be the highest sub-concept below this set, but it can also be a member of this set. The *supremum* of a set of concepts is the least concept of the context that is greater than or equal to all the elements of this set. It can also be said to be the nearest predecessor of this set.

□



**Figure 7.1** Lattice of “Some Life and Water” context.

FCA forces a hierarchical ordering of concepts. The data analysis techniques of this study are derived from this framework. Formal concept lattices and formal concept analysis are synonyms.

A line diagram of the lattice of the *some life and water context* is shown in Figure 7.1. The ordering scheme of the concepts is suggested by traversing the graph downwards from top to bottom, which corresponds to greatest and least respectively. The line diagram is a visual representation of a lattice and has many other uses, though it has not been used in the study. Its definition has been given as background to FCA, concepts and the prime operation are used however.

## 7.2 FCA Implications

### 7.2.1 Basic Definitions of FCA Implications

An implication has the following format  $A \rightarrow B$ ,  $A$  and  $B$  are sets of attributes of the context under consideration. For example, in the context of tuberculosis data that was collected in the study, the following is an implication that was found:  $\{Weight\ Loss, Night\ Sweat\} \rightarrow \{Tuberculosis\}$ . It can be loosely interpreted to mean that *weight loss* and *night sweat* implies the presence of *tuberculosis*. In the study it is said that the method of FCA implications were used to generate rules (implications). It can also perhaps be strictly argued that FCA implications are not a method, but in the study they have been treated as a method of FCA. The way that the term FCA implications is used should be clear whether it is the *method FCA implications* referred to, or whether it is a rule (*an implication*). Next the definition of an implication to hold in a context is given.

**Definition 7.6** (Ganter & Wille, 1999)

An implication  $A \rightarrow B$  *holds* in a context  $K = (G, M, I)$  if every object's intent respects  $A \rightarrow B$ . A subset of attributes  $T \subseteq M$  of a context *respects* an implication  $A \rightarrow B$  under the following conditions:

- If  $A \subseteq T$  then  $B \subseteq T$ .
- If  $A$  is not a subset of  $T$ , then  $T$  always respects the implication.

$T$  is called a *model* of  $A \rightarrow B$ .

□

An implication  $A \rightarrow B$  holds in a context if, every object that has attributes  $A$  also has  $B$ . That is,  $A$  always appears with  $B$ . In the case of the CDSS, all examples (study subject patients) must respect an implication (it must not be contradicted for it to be valid). The models of an implication are those sets of attributes where if  $A$  appears then  $B$  also appears. Note:  $B$  does not have to appear with  $A$ . A model respects an implication and an implication holds against a model. The following proposition suggests a method to check that an implication holds in a context.

**Proposition 7.2** (Ganter & Wille, 1999)

An implication  $A \rightarrow B$  holds in a context  $K = (G, M, I)$  iff  $B \subseteq A''$ , this is equivalent to saying that  $A' \subseteq B'$ . It then automatically holds in the set of all intents. ( $A''$  is called the closure of  $A$ )

□

A *closure system* is an important mathematical definition which is used in FCA. There is a proposition which is not made in the dissertation, but which links a closure system to a formal concept lattice. In this chapter closure systems are used in some propositions and definitions that are made.

**Definition 7.7** (Ganter & Wille, 1999)

A closure system on a set  $M$  is a set  $C \subseteq \mathcal{B}(M)$  which satisfies the following conditions:

- $M \in C$  and,
- if  $D \subseteq C$ , then  $\bigcap D \in C$ .

□

The following proposition identifies the intents of a context as the set which exactly respects the implications of a context, i.e. they are models of the implications of a context.

**Proposition 7.3** (Ganter & Wille, 1999)

If  $L$  is the set of implications of a context,  $K = (G, M, I)$  the set of models of  $L$  are given by  $\text{Mod } L = \{T \subseteq M \mid T \text{ respects } L\}$  and are a closure system. And if  $L$  is the set of all implications of a context then  $\text{Mod } L$  is exactly the system of all concept intents.

Proof:

- Firstly it is shown that  $\text{Mod } L$  is a closure system. An implication  $imp: A \rightarrow B$  is chosen. It can be asked: does  $M$  respect  $imp$ ? Yes, it does, because  $A \subseteq M$  and  $B \subseteq M$ . Hence  $M$  is a model of  $L$ , that is:  $M \in \text{Mod } L$ . Secondly, it must be shown that if  $D \subseteq \text{Mod } L$  then  $\bigcap D \subseteq \text{Mod } L$ . Choose  $S, T \in \text{Mod } L$  but  $S \cap T \notin \text{Mod } L$ . This happens when  $A \subseteq S \cap T$  but  $B \not\subseteq S \cap T$ . It has been argued that  $B \subseteq S$  and  $B \subseteq T$  because  $T$  and  $S$  respect  $A \rightarrow B$ . This is a contradiction because if  $B \subseteq S$  and  $B \subseteq T$  it must also be a subset of the intersection of  $S$  and  $T$ , that is,  $B \subseteq S \cap T$ . Hence  $S \cap T \subseteq \text{Mod } L$ . Therefore,  $\text{Mod } L$  is a closure system.
- It must now be demonstrated that  $\text{Mod } L$  is exactly the system of intents of the context if  $L$  is the set of implications of a context. Choose implication  $A \rightarrow B$  of the context, also choose  $C$  an intent which does not respect  $A \rightarrow B$ . This happens when  $A \subseteq C$  and  $B \not\subseteq C$ . Since  $C$  is a closure which contains  $A$ , then  $A'' \subseteq C$ . But  $B \subseteq A''$  because  $A \rightarrow B$  is an implication which holds in the context. This is a contradiction, so  $B \subseteq C$  hence  $C$  is a member of  $\text{Mod } L$ . Therefore all intents are members of  $\text{Mod } L$ . It must now be shown that only intents respect  $L$ . Choose  $C \subseteq M$  not an intent, at least one implication which  $C$  does not respect must be found; choose  $C \rightarrow C''$ . This is a valid implication of the context, hence a member of  $L$ , check proposition above on conditions for an implication to hold in a context. The premise of  $C \rightarrow C''$ ,  $C \subseteq C$ , but the conclusion  $C''$  is not a subset of  $C$ , hence  $C$  is not a model of  $L$ .

□

The above proof demonstrates some of the techniques and ideas used in FCA manipulations. A *closure system* was used in the first bullet.

## 7.2.2 The Duquenne-Guigues Basis

The implications of a context can be very many. They may not be good (easy) to work with. A set of implications called the Duquenne-Guigues basis has been defined. It was named as a result of work by Vincent Duquenne and J.L. Guigues (Guigues & Duquenne, 1986). It can be shown that all other implications of a context can be derived from this set. Further it can be shown that this set is *sound*, *complete* and *non-redundant*. *Sound* means that each implication in the Duquenne-Guigues basis holds in the context. *Complete* means that each implication in the context can be derived from the Duquenne-Guigues basis. By *non-redundant* it is meant that no implication in the Duquenne-Guigues basis follows from other implications within the set. What is meant by an implication to follow from another is defined next.

**Definition 7.8** (Ganter & Wille, 1999)

An implication  $A \rightarrow B$  is said to *follow* from a set  $L$  of implications in  $M$  if each subset of  $M$  respecting  $L$  also respects  $A \rightarrow B$ .

□

The following proposition suggests a method that can be used to derive other implications from an existing implication set. The three rules are also called Armstrong Rules (Silberschatz, Henry, & Sudarshan, 2002). They can be used to derive other implications of the context from the Duquenne-Guigues basis.

**Proposition 7.4** ((Ganter & Wille, 1999), (Stumme, 2003))

A set  $L$  of implications in  $M$  is closed iff the following conditions are satisfied for all  $W, X, Y, Z \subseteq M$ :

- $X \rightarrow X \in L$ .
- $X \rightarrow Y \in L$  then  $X \cup Z \rightarrow Y \in L$ .
- $X \rightarrow Y \in L$  and  $Y \cup Z \rightarrow W$ , then  $X \cup Z \rightarrow W \in L$ .

□

The following definition is of the pseudo-closed set which is used in the definition of the Duquenne-Guigues basis.

**Definition 7.9** (Ganter & Wille, 1999)

A pseudo-closed set  $P \subseteq M$  is defined recursively as follows.

- $P \neq P''$ , and
- if  $Q \subset P$  is a pseudo-closed proper subset of  $P$ , then  $Q'' \subseteq P$ .

□

The definition of the Duquenne-Guigues basis is given by the following proposition.

**Proposition 7.5** ((Guigues & Duquenne, 1986) (Ganter & Wille, 1999))

The Duquenne-Guigues basis is defined as follows:

$$\{P \rightarrow P'' \mid P \text{ is pseudo-closed}\}$$

□

The proof that the Duquenne-Guigues basis is sound, complete and non-redundant can be found in (Ganter & Wille, 1999). Proposition 7.5 uses pseudo-closed sets to find the Duquenne-Guigues basis. Finding pseudo-closed sets is not an easy task, but there is a method for deriving the Duquenne-Guigues basis, the reference ((Ganter & Wille, 1999), (Stumme, 2003)) explains.

### 7.2.3 Concept Explorer: Automated Tool for FCA Implications

In the project a tool called *Concept Explorer* was used to generate implications from the contexts patient data. It allows entering of objects using a cross table. Its interface is graphical user interface (GUI) based and it also has a user manual. It is a free tool that can be found on *Sourceforge* on the internet at <http://sourceforge.net/projects/conexp>, it is less than two megabytes in size. It is Java-based and at least JRE Version 5 is needed. It is possible to draw a line diagram, which can be stored as a JGP image that can be printed. It can also generate association rules which is another FCA technique but which was not considered in the study.

## 7.3 The JSM Method

### 7.3.1 Introduction to the JSM Method

The JSM method is named in honour of nineteenth century English philosopher, John Stuart Mill who proposed schemes of inductive reasoning. J. S. Mill suggested that common effects are likely due to common causes. The JSM method was originally formulated in the nineteen eighties in terms of predicate logic by V.K. Finn. The FCA formulation of this method is described by Ganter and Kuznetsov (2000).

The method employs *positive examples* and *negative examples* for a *target attribute*. Positive examples are objects which have the target attribute and negative examples are objects which do not have the target attribute. For example in the tuberculosis dataset each patient was deemed an example (object). Tuberculosis was the target attribute. Other attributes such as cough and weight loss, are called *structural attributes*. Those patients who had tuberculosis-like symptoms and who were diagnosed with tuberculosis were considered positive examples. Those patients who had tuberculosis-like symptoms but were not diagnosed with tuberculosis were considered negative examples. They may, for example, have had acute bronchitis.

By collecting negative and positive examples the JSM method hopes to generate hypotheses which can be used to predict the target attribute. In the study the JSM method was used to generate hypotheses which could be used to predict likelihood of illness.

### 7.3.2 Definition of a Hypothesis

According to the JSM method the context of the set of all objects  $G$  is split into three sub-contexts.

**Definition 7.10** (Ganter & Kuznetsov, 2000)

There is  $G_+$ , the positive examples,  $G_-$ , the negative examples, and  $G_t$ , the undetermined examples. The undetermined examples are not known whether they have or do not have  $w$  - the target attribute. They will be classified using the hypotheses. The three resulting contexts are;  $K_+$  of positive examples, and  $K_-$  of negative examples, and  $K_t$  of undetermined examples.

□

**Definition 7.11** (Ganter & Kuznetsov, 2000)

A *positive hypothesis* for  $w$  is an intent of  $K_+$  that is not contained in any intent of an object of  $K_-$ . It is an intent (closed set of attributes) found only in positive example. The set of attributes chosen are not just any combination but are intents. Why only intents are chosen will be explained later. Likewise, a *negative hypothesis* is an intent of  $K_-$  that is not found in any object of  $K_+$ .

□

A hypothesis is the premise of an implication or rule. Hence in the study it was said that the JSM method was used to generate rules. Another term used in the study is that a hypothesis corresponds to a rule (an implication).

### 7.3.3 Example

The following example taken from Blinova *et al.* (2003) is provided. Consider a context which is as follows:

$M = \{a, b, c, d, e\}$ , the attributes of the context.

$G_+ = \{g_1, g_2, g_3, g_4\}$ , the positive examples.

$G_- = \{g_5, g_6, g_7\}$ , the negative examples.

$G_t = \{g_8, g_9\}$ , the undetermined example.

The incidence relations are given by the following cross-table.

**Table 7.2 Example context from Blinova et al (2003).**

	a	b	c	d	e
g <sub>1</sub>	X	X	X		
g <sub>2</sub>	X	X		X	
g <sub>3</sub>	X	X			X
g <sub>4</sub>	X		X		X
g <sub>5</sub>	X		X	X	
g <sub>6</sub>		X	X	X	
g <sub>7</sub>	X			X	X
g <sub>8</sub>	X	X	X		X
g <sub>9</sub>			X	X	X

The object intents (attributes of objects) are as follows ( $g^+$ ,  $g^-$  and  $g^t$  are used denote the priming operation for an object of the positive, negative and undetermined examples' contexts respectively):

$$I_+ : g_1^+ = \{a, b, c\}, g_2^+ = \{a, b, d\}, g_3^+ = \{a, b, e\}, g_4^+ = \{a, c, e\}$$

$$I_- : g_5^- = \{a, c, d\}, g_6^- = \{b, c, d\}, g_7^- = \{a, d, e\}$$

$$I_t : g_8^t = \{a, b, c, e\}, g_9^t = \{c, d, e\},$$

Besides the intents of examples (objects) the following intents have also been found:

- Intents of positive examples context:

{a, b} its corresponding extent is {g<sub>1</sub>, g<sub>2</sub>, g<sub>3</sub>}, {a, e} extent {g<sub>3</sub>, g<sub>4</sub>}, {a, c} extent {g<sub>1</sub>, g<sub>4</sub>}, {a} extent {g<sub>1</sub>, g<sub>2</sub>, g<sub>3</sub>, g<sub>4</sub>}.

- Intents of negative examples context:

{c, d} extent {g<sub>5</sub>, g<sub>6</sub>}, {a, d} extent {g<sub>5</sub>, g<sub>7</sub>}, {d} extent {g<sub>5</sub>, g<sub>6</sub>, g<sub>7</sub>}.

The following can be found from the context:

- Positive hypotheses: {a, b} only. For example, positive intent {a, e} is not a hypothesis because it is contained in negative example g<sub>7</sub><sup>-</sup>.

- Negative hypotheses:  $\{c, d\}$  only. For example, negative intent  $\{a, d\}$  is not a hypothesis because it is contained in positive example  $g_2^+$ .
- Classification of undetermined examples:  $g_8$  is classified positive because it contains positive hypothesis  $\{a, b\}$ .  $g_9$  is classified negative because it contains the negative hypothesis  $\{c, b\}$ . It is possible for an undetermined example to contain both a negative and a positive hypothesis in that instance it is classified as being contradictory.

### 7.3.4 Choice of Intents as Hypotheses

Hypotheses are chosen among intents of a context because it is assumed that hypotheses are closed with respect to implications to other structural attributes. This means that hypotheses respect all the implications of the context of negative and positive examples. It was shown above Proposition 7.3 that the set which respects the implications are exactly the intents of the context.

But why are other subsets of attributes not selected? For example: some subsets of the intents which do not appear in negative examples but which always appear with the intents that are hypotheses. This is because hypotheses follow a cautious strategy. For example, say a positive hypothesis is given by the intent,  $C \cup D$ , since it has been observed that  $C$  always appears with  $D$  then it can be said that  $C \cup D$  is responsible for the target rather than  $C$  only. If  $D$  is thought to be superfluous then an example must be observed where given  $C$ ,  $D$  does not hold, but  $w$  does. This approach has been called conservative in the study. This is explained further in Chapter 8.

The following proposition suggests a way of finding hypotheses using pseudo-intents ((Ganter & Wille, 1999), (Stumme, 2003)).

#### Proposition 7.6

If  $P$  is a pseudo-intent of  $K_0$  such that  $w \notin P$  but  $w \in P_0''$ , then  $P_0'' \setminus \{w\}$  is a positive hypothesis.  $P_0''$  is an intent of the context of positive and negative example which does not include the target attribute and  $K_0$  is the context of positive and negative examples which does not include the target attribute.

□

### 7.3.5 QUDA: Automated Tool for JSM Hypotheses

The automated tool that was used to generate hypotheses is called *QUDA*. It has a GUI and a distribution of the software program can be found at <http://kirk.intellektik.informatik.tu-darmstadt.de/~quda/QDA.zip>. *QUDA*, like Concept Explorer requires the Java runtime to work. It supports a number of other data analysis techniques and data input techniques. The data input technique used in the project was the comma separated values (CSV) type.

## 7.4 Examples of Rules by JSM Hypotheses and FCA Implications

A few examples of rules generated with the FCA methods are presented. It begins with the JSM method.

### 7.4.1 Rules by JSM Method

In the example given in section 7.3.3 it was found that the set  $\{a, b\}$  is a positive hypothesis. It is an intent found only in positive examples for some target attribute  $w$ . This hypothesis corresponds to the following rule:  $\{a, b\} \rightarrow \{w\}$ .

Using the tuberculosis dataset as another example, the set  $\{WL, CP\}$  was found to be a positive hypothesis for the target attribute TB. WL is short for weight loss, and CP for chest pain, and TB for tuberculosis. This hypothesis corresponds to the following rule  $\{WL, CP\} \rightarrow \{TB\}$ . The rule can be interpreted as follows: if a person has symptoms weight loss and chest pain, then tuberculosis can be suspected.

### 7.4.2 Rules by FCA Implications

```
17 < 10 > WL NS ==> TB;
```

The above implication is taken from the tuberculosis results, it was generated using Concept Explorer. It is implication number 17, the number between angled brackets is the support for the implication and in this case it is 10. The support is the number of patients (examples in the

context of patient data) where the combination of symptoms of the implication is observed. The implication is formatted according to the style of Concept Explorer. It can be rewritten simply as the rule  $\{WL, NS\} \rightarrow \{TB\}$ . This rule states that if a person has the symptoms WL (weight loss) and NS (night sweat), then TB (tuberculosis) can be suspected. Another example which is different to this one is the following.

```
16 < 6 > PC SP NS ==> WL TB;
```

From this implication the following implication can be derived:  $\{PC, SP, NS\} \rightarrow \{TB\}$ . This is valid by the definition of implications which states that the premise of a rule also implies each member of the conclusion. This rule concludes on tuberculosis, and was included in the study.

The implications used in the study were limited to the Duquenne-Guigues basis. From these, only those implications that have the target attribute on the right hand side were selected. For example, for tuberculosis only those implications which have TB on the right hand side were considered.

## 7.5 Conclusion

In this chapter the theoretical foundations of the FCA methods were described. Automated tools that were used were also presented. They are Concept Explorer for FCA implications and QUDA for JSM hypotheses. It was shown how implications were used to create new rules. Next is data analysis.

## Chapter 8

### Data Analysis

This chapter presents the results of data analysis. The results are for the illnesses for which data was collected, these are: tuberculosis, bronchitis, and hypertension. Malaria and peptic ulcers have not been included due to data collection limitations. Nevertheless, the results which are discussed in this chapter were found adequate to decide preliminarily the utility of FCA at evaluation of the CDSS.

For each illness there was a training dataset which was used to generate implications and JSM hypotheses, these are also called FCA generated rules. These rules that were generated were compared to the original experts' rules as follows: *rules subsumed*, *new rules* and *rules not accounted*. The new rules were then tested by measuring the performance of the CDSS at giving advice using the test dataset.

There was not enough data for bronchitis. Hence the bronchitis data was all used for the training dataset. As a result the CDSS performance for bronchitis was not done.

At the time of data collection there were no experts' rules for hypertension. However, the decision was made to go ahead, and data was collected. In this case the FCA methods were used to find out what rules could be generated from the patients' data. A training dataset and a test dataset were available, and the performance for the new rules was measured.

Summary tables for data analysis are presented in section 8.5. The detailed analysis appears in the preceding sections. The conclusion to the chapter contains a general discussion on the results, as well as the discussion of the experts on the results.

Before presenting the results, an explanation of how to interpret the results is given. This explains how to interpret a cross table which is how the data is formatted for the FCA methods. The notation of rules and implications is also revised. The following subheadings are used.

- 8.1 How to interpret the results.
- 8.2 Analysis of tuberculosis.

- 8.3 Analysis of bronchitis.
- 8.4 Analysis of hypertension.
- 8.5 Summary of data analysis.
- 8.6 Chapter conclusion.

## 8.1 How to Interpret Results

**Table 8.1 Example context, how to interpret results.**

	PC	SP	MC	BS	CS	WL	NS	NA	CP	SB	TC	TN	TB
Obj9							X						X
Obj10	X	X	X			X	X	X	X			X	X
Obj11	X	X	X			X		X		X			

For each illness the data collected was formatted into a cross table. Table 8.1 is an example from the tuberculosis data. Rows are labeled *Obj1*, *Obj2*, and so on (*Obj* is short for object). Each row corresponds to a single study patient. For example, *Obj10* is subject patient number 10. *Obj10* (patient 10) presented with the following symptoms: PC (Persistent Cough), SP (Sputum Production), MC (Sputum produced is Muco-purulent), WL (Weight Loss), NS (Night Sweat), NA (No Appetite), CP (Chest Pain), and TN (Tiredness). The presence of a symptom – an attribute – is denoted by a cross. The last column records the diagnosis by the treating doctor. The diagnosis is also called the target attribute. In the case of *Obj10* the diagnosis is TB (tuberculosis). *Obj10* is classified as a positive example. In the case of *Obj11* (Subject Patient number 11) the diagnosis is not TB, and the patient has symptoms: PC, SP, MC, WL, NA and SB (Shortness of Breath). *Obj11* is classified a negative example. For each illness a key is given which explains the abbreviations used.

For each illness the rules given by the experts are listed as well as those generated by the FCA methods. An example of how to interpret a rule is given next.

1. PC SP BS WL → TB
2. PC SP BS CP NA → TB

Rule number 1 can be interpreted as follows: If a patient has PC (Persistent Cough), SP (Sputum Production), BS (Sputum produced is Bloody), WL (Weight Loss) then she may be suspected to

have TB (Tuberculosis). A rule has a left hand side, called a premise and a right hand side, called a conclusion. The premise and conclusion are separated by an arrow. The premise is the set of conditions which are needed in order for the conclusion of the rule to be applicable.

The set of implications generated by FCA can be interpreted likewise (an implication is equivalent to a rule). An example is given.

1 < 6 > SP  $\implies$  PC;  
17 < 10 > WL NS  $\implies$  TB;  
18 < 10 > WL CP  $\implies$  TB;

Implication number 17 states that WL (weight loss) and NS (night sweat) imply TB (Tuberculosis), according to the dataset which was collected. This was a useful finding which was taken back to the experts for consideration. Implication number 18 states that WL and CP also imply TB. Implication number 1 states that if a patient produces sputum (SP) then they may have a PC (Persistent Cough) – an implication which might have been expected *a priori*, but which is not considered particularly useful for the CDSS.

An implication is equivalent to a rule. It has a left hand side which is called the premise; and a right hand side, which is called the conclusion. The premise and conclusion are also separated by an arrow. The number between angled brackets is the support for the implication. It is the number of patients (objects) in the dataset where the implication is observed, that is, the number of patients where the combination of symptoms of the implication is found.

Only the Duquenne-Guigues basis of implications are listed in the results. This because this set is non-redundant, also refer to Chapter 7, Section 7.2.2. This means that no implication of the set can be derived from another within the set. It is also complete which means that all possible implications of the context can be derived from the Duquenne-Guigues basis of a context. An implication of the Duquenne-Guigues basis can have zero (0) support, and still be valid. It is valid in the sense that it does not contradict any example of the context. The context describes the patient data that was collected. Concept Explorer, which is the automated tool for implications used in the project, also lists those implications which have zero support. In the study implications which have zero support were not considered. Only implications which make positive predictions about the illness are of interest in the study. For example, in the tuberculosis dataset only the implications that have TB on the right hand side were considered.

JSM hypotheses are sets of attributes which can be used to predict the target attribute. They are actually premises of rules. In the dissertation a hypothesis is said to have a corresponding implication or rule. For example, from the hypertension data the following set of symptoms {WL, CP} was found to be a positive hypothesis for the target attribute TB. WL is short for weight loss and CP for chest pain and TB for tuberculosis. This hypothesis corresponds to the following (implication) rule  $\{WL, CP\} \rightarrow \{TB\}$ . The rule can be interpreted as follows: if a person has symptoms weight loss and chest pain, then it can be suspected that tuberculosis is present. JSM hypotheses are a subset of the implications of a context. That is, they always have corresponding implications whose premises are the same. Negative hypotheses were not considered in the study. This is because in the current phase of the project the possibility of deciding that an illness is absent is not considered.

## 8.2 Analysis of Tuberculosis

**Table 8.2 Tuberculosis (TB) symptoms.**

<u>Symptoms</u>
PERSISTENT COUGH ( <b>PC</b> )
SPUTUM PRODUCTION ( <b>SP</b> )
MUCO-PURULENT ( <b>MC</b> )
SPUTUM BLOODY ( <b>BS</b> )
CLEAR-SPUTUM ( <b>CS</b> )
WEIGHT LOSS ( <b>WL</b> )
EXTREME NIGHT SWEATS ( <b>NS</b> )
NO APPETITE ( <b>NA</b> )
CHEST PAIN ( <b>CP</b> )
SHORTNESS OF BREATH ( <b>SB</b> )
TUBERCULOSIS CONTACT ( <b>TC</b> )
TIREDNESS ( <b>TN</b> )

## 8.2.1 Experts' Rules for Tuberculosis

The rules for diagnosing tuberculosis that were initially elicited from the experts are the following:

1. PC SP BS WL → TB
2. PC SP BS NS → TB
3. PC WL NS → TB
4. PC SP BS TC → TB
5. PC SP BS CP NA → TB
6. PC SP BS SB → TB
7. PC WL CP SB → TB
8. PC SP BS CP TN → TB

## 8.2.2 Training Dataset for Tuberculosis

**Table 8.3 Training dataset for tuberculosis.**

Note: the diagnosis on the last column is provided by the treating doctor, and the symptoms by the patients via the CDSS questionnaire.

	PC	SP	MC	BS	CS	WL	NS	NA	CP	SB	TC	TN	TB
Obj1	X	X	X			X			X	X			X
Obj2	X	X	X			X	X	X	X				X
Obj3	X	X	X			X	X	X	X		X	X	X
Obj4	X					X		X		X		X	X
Obj5						X	X	X	X	X		X	X
Obj6	X					X	X	X	X	X	X	X	X
Obj7	X	X			X	X		X	X	X		X	X
Obj8						X	X					X	X
Obj9							X						X
Obj10	X	X	X			X	X	X	X			X	X
Obj11	X	X	X			X		X		X			
Obj12	X	X	X			X		X					
Obj13	X	X			X				X			X	
Obj14	X	X		X		X	X	X	X			X	X
Obj15									X				
Obj16	X	X	X			X	X	X	X		X	X	X
Obj17	X					X	X	X	X	X		X	X
Obj18	X	X			X	X	X					X	X

	PC	SP	MC	BS	CS	WL	NS	NA	CP	SB	TC	TN	TB
Obj19	X						X	X				X	
Obj20	X	X	X						X			X	
Obj21									X			X	

The tuberculosis dataset was split into two. Table 8.3, shows the training dataset, which was used to generate new rules. The test dataset, the second part of the data is shown in Table 8.6, it was used to test the new rules.

## 8.2.3 FCA Implications for Tuberculosis

### 8.2.3.1 Duquenne-Guigues Basis Implications

- 1 < 12 > SP  $\implies$  PC;
- 2 < 8 > MC  $\implies$  PC SP;
- 3 < 3 > CS  $\implies$  PC SP TN;
- 4 < 8 > PC SP NA  $\implies$  WL;
- 5 < 7 > SB  $\implies$  WL;
- 6 < 4 > PC SP MC TN  $\implies$  CP;
- 7 < 11 > PC TB  $\implies$  WL;
- 8 < 5 > PC SP MC WL TB  $\implies$  CP;
- 9 < 10 > NA TB  $\implies$  WL;
- 10 < 6 > PC SP WL NA TB  $\implies$  CP;
- 11 < 10 > CP TB  $\implies$  WL;
- 12 < 2 > PC SP WL SB TB  $\implies$  CP;
- 13 < 5 > WL NA SB TB  $\implies$  TN;
- 14 < 11 > TN TB  $\implies$  WL;
- 15 < 1 > BS  $\implies$  PC SP WL NS NA CP TN TB;
- 16 < 6 > PC SP NS  $\implies$  WL TB;
- 17 < 10 > WL NS  $\implies$  TB;
- 18 < 10 > WL CP  $\implies$  TB;
- 19 < 8 > NS CP  $\implies$  WL NA TB;
- 20 < 9 > NA CP  $\implies$  WL TB;
- 21 < 3 > TC  $\implies$  PC WL NS NA CP TN TB;
- 22 < 11 > WL TN  $\implies$  TB;
- 23 < 0 > PC SP MC CS CP TN  $\implies$  BS WL NS NA SB TC TB;
- 24 < 8 > WL NS NA TB  $\implies$  CP;
- 25 < 4 > PC SP MC WL NA CP TB  $\implies$  NS;
- 26 < 3 > WL NS SB TB  $\implies$  NA CP TN;

- 27 < 8 > WL CP TN TB  $\implies$  NA;  
 28 < 5 > WL SB TN TB  $\implies$  NA;  
 29 < 1 > PC SP WL NA CP SB TN TB  $\implies$  CS;  
 30 < 2 > PC SP WL NS NA CP TC TN TB  $\implies$  MC;  
 31 < 1 > PC SP CS WL NA CP TN TB  $\implies$  SB;  
 32 < 0 > PC SP CS WL NS NA CP SB TN TB  $\implies$  MC BS TC; and  
 33 < 0 > PC SP MC BS WL NS NA CP TN TB  $\implies$  CS SB TC.

Note: Only implications which conclude about TB are considered. These are the implications which have TB on the right hand side.

### 8.2.3.2 Experts' Rules Subsumed (Accounted)

- Experts' rule 2. PC SP BS NS  $\rightarrow$  TB has been subsumed as follows. Implication 16 is given by the following:  $\{PC, SP, NS\} \rightarrow \{WL, TB\}$ . From this implication, the following implication follows (can be derived):  $\{PC, SP, NS\} \rightarrow \{TB\}$ . From this implication, the following implication follows  $\{PC, SP, BS, NS\} \rightarrow \{TB\}$ , which corresponds to the experts' rule. Thus it can be said that experts' rule 2 is subsumed by an FCA generated rule, that is, the experts' rule corresponds to an FCA implication. From this case the following proposition can be made.

#### Proposition 8.1

If the premise of an implication which has the target attribute on the right hand side is a subset of the premise of an experts' rule, it can be said that the experts' rule follows from the implication and hence it can be considered subsumed. For example, in the above case, the premise of implication 16 is  $\{PC, SP, NS\}$ , which is a subset of the premise of experts' rule 2,  $\{PC, SP, BS, NS\}$ . Implication 16 also has the target attribute on its right hand side, which means it concludes about the illness. Hence experts' rule 2, is classified subsumed (or accounted for).

□

- Experts' rule 3. PC WL NS  $\rightarrow$  TB has been subsumed by implication 17. The premise of implication 17,  $\{WL, NS\}$ , is a subset of the premise of this experts' rule. And the right hand

side of the implication has the target attribute. So by Proposition 8.1 the experts' rule is classified subsumed by the FCA implication.

- The following experts' rules are subsumed by an implication 15.

1. PC SP BS WL  $\rightarrow$  TB

4. PC SP BS TC  $\rightarrow$  TB

5. PC SP BS CP NA  $\rightarrow$  TB

6. PC SP BS SB  $\rightarrow$  TB

8. PC SP BS CP TN  $\rightarrow$  TB

The premise of implication 15, {BS}, is a subset of the premise of each of these experts' rules. The right hand side of the implication has the target attribute. By Proposition 8.1 the experts' rules follow from the implication and thus can be classified as subsumed.

- Experts' rule 7. PC WL CP SB  $\rightarrow$  TB has been subsumed implication 18. The premise of implication 18, {WL, CP}, is a subset of the premise of this experts' rule. The right hand side of the implication has the target attribute. By Proposition 8.1 above the experts' rule is subsumed by the FCA implication.

Proposition 7.4 in Section 7.2.2 of Chapter 7 can be used to derive implications from.

For Tuberculosis, there were no implications of the Duquenne-Guigues basis which overlapped experts' rules exactly, though all experts' rules were subsumed. This observation was used in the in the conclusion to suggest that FCA implications found more *opportunistic* rules than did the experts. There are two terms which were used to characterize rules in the study. These are: *conservative* and *opportunistic*. They are explained further in the conclusion of this chapter. For now it can be said that conservative means a rule with a larger premise, that is, it needs more symptoms before concluding than an opportunistic rule. Perhaps there should have been a fourth category for comparison. In addition to *rules subsumed*, *new rules*, *rules not accounted for*, there should have been the category *non-redundant overlapping rules*. By non-redundant it is meant a rule that is of the Duquenne-Guigues basis. Concept explorer was used to generate the Duquenne-Guigues basis in the study.

### 8.2.3.3 New Rules

- A list of new rules corresponding to implications whose premises are not subsets of any premise of an existing experts' rule, and which of course have the target attribute on the right hand side are given first.

9. NS CP  $\rightarrow$  TB

10. NA CP  $\rightarrow$  TB

11. WL TN  $\rightarrow$  TB

New rule 9 is implication 19. New rule 10 is implication 20. New rule 11 is implication 22.

- The premises of the following implications are subsets of the premises of existing experts' rule. These implications have been used in the classification of experts' rules as subsumed. They can be described as more opportunistic versions of the experts' rules. They are opportunistic in the sense that they make diagnosis based on fewer symptoms than the experts' rules require.

12. BS  $\rightarrow$  TB

13. PC SP NS  $\rightarrow$  TB

14. WL NS  $\rightarrow$  TB

15. WL CP  $\rightarrow$  TB

16. TC  $\rightarrow$  TB

New rule 12 is implication 15. New rule 13 is implication 16. New rule 14 is implication 17.

New rule 15 is implication 18. New rule 16 is implication 21.

### 8.2.3.4 Rules Unaccounted

All experts' rules for tuberculosis were accounted for by FCA implications.

## 8.2.4 JSM Hypotheses for Tuberculosis

### 8.2.4.1 Hypotheses

1. {WL, CP}

2. {WL, TN}

3. {WL, NS}

### 8.2.4.2 Experts' Rules Subsumed

- Experts' rule 7. PC WL CP SB  $\rightarrow$  TB has been subsumed by hypothesis 1. Hypothesis 1  $\{WL, CP\}$ , corresponds to the following implication  $\{WL, CP\} \rightarrow \{TB\}$ . The premise of this implication is a subset of the premise of this experts' rule. The right hand side of the implication has the target attribute. By Proposition 8.1 the experts' rule is subsumed by the FCA implication.
- Experts' rule 3. PC WL NS  $\rightarrow$  TB has been subsumed by hypothesis 3. Hypothesis 3,  $\{WL, NS\}$ , corresponds to implication  $\{WL, NS\} \rightarrow \{TB\}$ . The premise of this implication is a subset of the premise of this experts' rule. The right hand side of the implication has the target attribute. By Proposition 8.1 the experts' rule is subsumed by the FCA implication.

### 8.2.4.3 New Rules

17. WL CP  $\rightarrow$  TB

18. WL TN  $\rightarrow$  TB

19. WL NS  $\rightarrow$  TB

New rule 17 corresponds to hypothesis 1. New rule 18 corresponds to hypothesis 2. New rule 19 corresponds to hypothesis 3.

JSM hypotheses are a subset of the implications of a context always. For example new rule 17 in this section is the same as new rule 15 of implications, Section 8.2.3.3. New rules 18 and 19 can also be found in the set generated by the implications.

#### 8.2.4.4 Rules Not Accounted

The following experts' rules have not been accounted for by JSM hypotheses for tuberculosis.

1. PC SP BS WL  $\rightarrow$  TB
2. PC SP BS NS  $\rightarrow$  TB
4. PC SP BS TC  $\rightarrow$  TB
5. PC SP BS CP NA  $\rightarrow$  TB
6. PC SP BS SB  $\rightarrow$  TB
8. PC SP BS CP TN  $\rightarrow$  TB

All these experts' rules have BS on the left-hand side. In the data, there is only one patient with this symptom, object 14. The intent of this object is {PC SP BS WL NS NA CP TN} which is also a hypothesis. It can be said that the JSM method does not have enough evidence to generalise that {BS}  $\rightarrow$  {TB}, it always observes BS with the other symptoms. That is, there is no example where BS is present and the other symptoms are not present. This hypotheses though is redundant because it is subsumed by hypothesis 1 , {WL CP}. Hence it was not included in the result.

#### 8.2.5 Summary of FCA Analysis for Tuberculosis

A comparison of the experts' rules with the rules generated by the FCA methods is made in this section. The comparison is done according to the criteria of the study design: *rules subsumed*, *new rules*, *rules not accounted for*. Table 8.4 is of rules subsumed and rules not accounted. And Table 8.5 is of the new rules.

Table 8.4 can be interpreted as follows. A value of 1 means that the experts' rule has been accounted for (subsumed) by the technique of the column. A value of 0, means that the experts' rule has not been accounted for by the technique. For example, the first row is experts' rule: PC SP BS WL  $\rightarrow$  TB . The value for FCA implications is 1, which means it can be classified subsumed using an implication of the tuberculosis dataset. For the JSM hypotheses the value is 0, which means that the experts' rule cannot be classified subsumed using a JSM hypothesis, that is, it is unaccounted for by the JSM hypotheses of the dataset.

**Table 8.4 Experts' rules subsumed, tuberculosis.**

<u>Experts' Rules</u>	<u>FCA Implications</u>	<u>JSM Hypothesis</u>
1. PC SP BS WL → TB	1	0
2. PC SP BS NS → TB	1	0
3. PC WL NS → TB	1	0
4. PC SP BS TC → TB	1	1
5. PC SP BS CP NA → TB	1	0
6. PC SP BS SB → TB	1	0
7. PC WL CP SB → TB	1	1
8. PC SP BS CP TN → TB	1	0

Table 8.5, of new rules, can be interpreted as follows. A value of 1, means that an FCA methods' rule has been generated by the technique of the column. A value of 0 means that rule cannot be derived by the technique. For the column "accepted by experts", 1 means that the rule has been accepted by experts, and 0 means that it was not accepted. The support for each rule is also given. For example new rule 9 is supported by 8 patients, this means that it was observed in 8 patients.

**Table 8.5 New rules, tuberculosis.**

<u>New Rules</u>	<u>FCA Implications</u>	<u>JSM Hypothesis</u>	<u>Accepted by Experts</u>	<u>Support</u>
9. NS CP → TB	1	0	1	8
10. NA CP → TB	1	0	1	9
11. WL TN → TB	1	1	1	11
12. BS → TB	1	0	1	1
13. PC SP NS → TB	1	0	1	6
14. WL NS → TB	1	1	1	10
15. WL CP → TB	1	1	1	10
16. TC → TB	1	0	1	3

## 8.2.6 CDSS Performance for Tuberculosis

Table 8.6 is the test dataset for tuberculosis it was used for running the CDSS performance at giving advice. Table 8.7 is the performance of the CDSS with the new rules included, against the performance of the CDSS with the original experts' rules only.

**Table 8.6 Test dataset for tuberculosis.**

	PC	SP	MC	BS	CS	WL	NS	NA	CP	SB	TC	TN	TB
Obj22	X								X			X	
Obj23	X	X	X			X	X		X	X		X	X
Obj24	X	X	X			X	X	X	X		X	X	X
Obj25	X	X	X			X		X	X	X	X	X	X
Obj26	X	X	X			X	X	X			X	X	X
Obj27	X	X	X			X	X		X	X	X	X	X
Obj28	X	X			X	X	X	X				X	X
Obj29	X					X	X		X			X	X
Obj30						X							X
Obj31						X	X	X					X

The conclusion (diagnosis) of the CDSS, Table 8.7, can be interpreted as follows. It either predicts presence of an illness, or it fails to predict illness. When it fails to predict illness it could be that the patient has another illness. Or the patient does not have a serious illness. Or it could be that it has failed to find the illness though it was present. But as stated it does not include rules which make negative predictions, so it cannot exclude illness. In its advice it will tell the user to go to the clinic for another examination, when it fails to predict illness. Later it may include negative predictions for minor illnesses like sore throat, in the absence of serious symptoms. But it was still considered to be beneficial in its present form. This is because it was suggested that many users of the CDSS may actually have a problem, but have not yet gone to seek help. Its role would be to encourage them to seek help.

The CDSS performance measures the number of times that it has the same diagnosis as the treating doctor. For example if out of ten cases it has the same diagnosis as the treating doctor seven times then the performance of the CDSS is rated at seventy percent (70%). If the CDSS and the doctor do not diagnose the illness under consideration then that is considered a match.

For example for object 22, the CDSS and the treating doctor did not diagnose tuberculosis, and this was considered a match.

**Table 8.7 Performance of CDSS for tuberculosis.**

<u>Patient</u>	<u>CDSS Assessment</u> <i>(New Rules Included)</i>	<u>Treating Doctor's</u> <u>Assessment</u>	<u>CDSS Assessment</u> <i>(Original</i> <i>Experts' Rules)</i>
Obj22	-	-	-
Obj23	Tuberculosis	Tuberculosis	Tuberculosis
Obj24	Tuberculosis	Tuberculosis	Tuberculosis
Obj25	Tuberculosis	Tuberculosis	Tuberculosis
Obj26	Tuberculosis	Tuberculosis	Tuberculosis
Obj27	Tuberculosis	Tuberculosis	Tuberculosis
Obj28	Tuberculosis	Tuberculosis	Tuberculosis
Obj29	Tuberculosis	Tuberculosis	Tuberculosis
Obj30	-	Tuberculosis	-
Obj31	Tuberculosis	Tuberculosis	-
<b>Performance</b>	90%		80%

### 8.3 Analysis of Chronic Bronchitis

In this section bronchitis is used as short for chronic bronchitis.

**Table 8.8 Chronic Bronchitis (CBr) symptoms.**

<b><u>Symptoms</u></b>
PERSISTENT COUGH ( <b>PC</b> )
SPUTUM PRODUCTION ( <b>SP</b> )
SPUTUM MUCO-PURULENT ( <b>MP</b> )
CHEST TIGHTNESS ( <b>CT</b> )
SHORTNESS OF BREATH ( <b>SB</b> )
WHEEZING CHEST ( <b>WC</b> )
SMOKING ( <b>SM</b> )

#### 8.3.1 Experts' Rules for Bronchitis

1. PC SM SP CT SB → CBr
2. PC SM SP WC → CBr
3. PC SM SP MC → CBr

#### 8.3.2 Training Dataset for Bronchitis

The bronchitis dataset was collected by initially enquiring whether or not a patient smoked. Only smokers were selected for further questioning. Within the person-power and time limitations of this early study, the best that could be achieved was the collection of data for seven patients, as reflected by Table 8.9. There is only one positive example in the dataset of chronic bronchitis, namely Obj4. Because of this limited dataset, it was not possible to follow the same methodology as for tuberculosis as outlined above – i.e. it was not feasible to split the data into a training set and a test dataset.

**Table 8.9 Training dataset for bronchitis.**

	PC	SP	MC	CS	CT	SB	WC	SM	CBr
Obj1	X	X	X					X	
Obj2	X							X	
Obj3	X	X		X	X			X	
Obj4	X	X	X		X	X	X	X	X
Obj5	X	X		X	X		X	X	
Obj6								X	
Obj7								X	

### 8.3.3 FCA Implications for Bronchitis

#### 8.3.3.1 Duquenne-Guigues Basis Implications

- 1  $\langle 7 \rangle \{ \} \implies SM$ ; (This implication simply reflects the fact that only smokers were used in this dataset.)
- 2  $\langle 4 \rangle SM \implies PC$ ;
- 3  $\langle 1 \rangle SM \ CBr \implies PC \ SP \ MC \ CT \ SB \ WC$ ;
- 4  $\langle 2 \rangle WC \ SM \implies PC \ SP \ CT$ ;
- 5  $\langle 1 \rangle SB \ SM \implies PC \ SP \ MC \ CT \ WC \ CBr$ ;
- 6  $\langle 3 \rangle CT \ SM \implies PC \ SP$ ;
- 7  $\langle 2 \rangle CS \ SM \implies PC \ SP \ CT$ ;
- 8  $\langle 2 \rangle MC \ SM \implies PC \ SP$ ;
- 9  $\langle 4 \rangle SP \ SM \implies PC$ ;
- 10  $\langle 2 \rangle PC \ SP \ CT \ SM \implies CS$ ; and
- 11  $\langle 1 \rangle PC \ SP \ MC \ CT \ SM \implies SB \ WC \ CBr$ .

#### 8.3.3.2 Experts' Rules Subsumed

- Experts' rule number 1.  $PC \ SM \ SP \ CT \ SB \rightarrow CBr$  has been subsumed by implication 5. The premise of implication 5,  $\{SB, SM\}$ , is a subset of the premise of this experts' rule. The right hand side of the implication has the target attribute, CBr. By Proposition 8.1 above the experts' rule is subsumed by the FCA implication.

#### 8.3.3.3 New Rules

- New rule 4.  $PC \ SP \ MC \ CT \ SM \rightarrow CBr$ . This rule may be considered a more conservative version of experts' rule 3. It is implication 11.

- New rule 5. SB SM → CBr, this is implication 5. It is a more opportunistic version of experts' rule 1.

#### 8.3.3.4 Rules Not Accounted

- Experts' rule 2. PC SM SP WC → CBr .
- Experts' rule 3. PC SM SP MC → CBr .

### 8.3.4 JSM Hypotheses for Bronchitis

#### 8.3.4.1 Hypotheses

1. {PC, SP, MC, CT, SB, WC, SM}

#### 8.3.4.2 Experts' Rules Subsumed

None of the experts' rules correspond to the JSM hypothesis that was found. There is only one positive example in our data. Hence, the only positive hypothesis is the intent of this example, which is not a subset of the premise of any of the experts' rules..

#### 8.3.4.3 New Rules

- The following implication 4. PC SP MC CT SB WC SM → CBr<sub>2</sub> corresponding to the hypothesis may be considered. This rule may be considered a more conservative version of experts' rule number 3. Recall that for tuberculosis all the new rules generated by FCA were more opportunistic. This rule however was not included in the results because it is subsumed by new rule 5 of the implications, and hence it is redundant.

#### 8.3.4.4 Rules Not Accounted

All the experts' rules were not accounted for by the JSM hypotheses for bronchitis.

### 8.3.5 Summary of FCA Analysis for Bronchitis

**Table 8.10 Experts' rules subsumed, bronchitis**

<u>Experts' Rules</u>	<u>FCA Implications</u>	<u>JSM Hypothesis</u>
PC SM SP CT SB → CBr	1	0
PC SM SP WC → CBr	0	0
PC SM SP MC → CBr	0	0

**Table 8.11 New rules, bronchitis.**

<u>New Rules</u>	<u>FCA Implications</u>	<u>JSM Hypothesis</u>	<u>Accepted by Experts</u>	<u>Support</u>
4. PC SP MC CT SM → CBr	1	0	1	1
5. SB SM → CBr	1	0	1	1

## 8.4 Analysis of Hypertension

At the time of data collection there were no experts' rules for hypertension. However, the decision was made to go ahead and data was collected and FCA methods were used to find out rules that could be generated from the data. These were then tested using the test dataset.

**Table 8.12 Hypertension (HP) symptoms.**

<b><u>Symptoms</u></b>
PERSISTENT OCCIPITAL HEADACHE ( <b>OH</b> )
BLURRED-VISION ( <b>BV</b> )
DIZZINESS ( <b>DZ</b> )
NAUSEA ( <b>NU</b> )
NOSE BLEED ( <b>NB</b> )
OVERWEIGHT ( <b>OW</b> )
FAMILY HYPERTENSION ( <b>FH</b> )
LIFESTYLE RISK ( <b>LR</b> )

### 8.4.1 Training Dataset for Hypertension

**Table 8.13 Training dataset for hypertension.**

	OH	BV	DZ	NU	NB	OW	FH	LR	HP
Obj1	X	X	X	X					X
Obj2	X	X	X		X			X	X
Obj3	X		X			X	X	X	X
Obj4	X		X				X	X	X
Obj5	X		X			X	X	X	X
Obj6	X	X	X	X	X			X	X
Obj7	X		X				X	X	X

	OH	BV	DZ	NU	NB	OW	FH	LR	HP
Obj8	X	X	X	X		X		X	X
Obj9			X					X	X
Obj10			X		X			X	X
Obj11	X	X					X	X	

## 8.4.2 FCA Implications for Hypertension

### 8.4.2.1 Duquenne-Guigues Basis Implications

- 1  $\langle 10 \rangle \text{HP} \implies \text{DZ}$ ;
- 2  $\langle 5 \rangle \text{FH} \implies \text{OH LR}$ ;
- 3  $\langle 3 \rangle \text{OW} \implies \text{OH DZ LR HP}$ ;
- 4  $\langle 3 \rangle \text{NB} \implies \text{DZ LR HP}$ ;
- 5  $\langle 3 \rangle \text{NU} \implies \text{OH BV DZ HP}$ ;
- 6  $\langle 10 \rangle \text{DZ} \implies \text{HP}$ ;
- 7  $\langle 5 \rangle \text{BV} \implies \text{OH}$ ;
- 8  $\langle 2 \rangle \text{OH DZ NB LR HP} \implies \text{BV}$ ;
- 9  $\langle 0 \rangle \text{OH BV DZ FH LR HP} \implies \text{NU NB OW}$ ;
- 10  $\langle 1 \rangle \text{OH BV DZ OW LR HP} \implies \text{NU}$ ;
- 11  $\langle 0 \rangle \text{OH BV DZ NU NB OW LR HP} \implies \text{FH}$ ;

### 8.4.2.2 New Rules

1.  $\text{OW} \rightarrow \text{HP}$ , this is implication 3.
2.  $\text{NB} \rightarrow \text{HP}$ , this is implication 4.
3.  $\text{NU} \rightarrow \text{HP}$ , this is implication 5.
4.  $\text{DZ} \rightarrow \text{HP}$ , this is implication 6

## 8.4.3 JSM Hypotheses for Hypertension

### 8.4.3.1 Hypotheses

1.  $\{\text{DZ}\}$

### 8.4.3.2 New Rules

4. DZ → HP , corresponds to the hypothesis.

### 8.4.4 Summary of FCA Analysis for Hypertension

**Table 8.14 New rules, hypertension.**

<u>New Rules</u>	<u>FCA Implications</u>	<u>JSM Hypothesis</u>	<u>Accepted by Experts</u>	<u>Support</u>
1. OW → HP	1	0	1	3
2. NB → HP	1	0	1	3
3. NU → HP	1	0	1	3
4. DZ → HP	1	1	1	10

### 8.4.5 CDSS Performance for Hypertension

At the time of data collection there were no experts' rules for hypertension. As a result the performance was measured for the new rules only. The aim was to check if the performance meets the target of 70% that was set initially.

**Table 8.15 Test dataset for hypertension.**

	OH	BV	DZ	NU	NB	OW	FH	LR	HP
Obj12	X				X	X	X	X	X
Obj13	X	X	X					X	X
Obj14						X	X		X
Obj15	X	X	X				X	X	X
Obj16	X		X	X					X
Obj17	X	X	X	X		X	X	X	X
Obj18			X			X	X		X
Obj19									X
Obj20	X		X						
Obj121	X	X					X	X	X

**Table 8.16 Performance of CDSS, hypertension.**

<u>Patient</u>	<u>CDSS Assessment</u>	<u>Treating Doctors' Assessment</u>
Obj12	Hypertension	Hypertension
Obj13	Hypertension	Hypertension
Obj14	Hypertension	Hypertension
Obj15	Hypertension	Hypertension
Obj16	Hypertension	Hypertension
Obj17	Hypertension	Hypertension
Obj18	Hypertension	Hypertension
Obj19	-	Hypertension
Obj20	Hypertension	-
Obj121	-	Hypertension
<b>Performance</b>	70%	

## 8.5 Summary of Data Analysis

Table 8.17 is a summary of the comparison between the experts' rules and FCA implications. Table 8.18 is a summary of the comparison between experts' rules and JSM hypotheses. The tables check how many experts' rules were subsumed by each technique, and how many rules were not accounted for, and how many new rules were generated.

**Table 8.17 Summary of comparison of experts' rules and FCA implications.**

<u>Item</u>	<u>TB</u> <i>8 experts' rules.</i>	<u>CBr</u> <i>3 experts' rules.</i>	<u>Hp</u> <i>No experts' rules.</i>	<u>Total</u> <i>11 experts' rules.</i>
Rules Confirmed	8	1	-	9
Rules Not Accounted	0	2	-	2
New Rules	8	2	4	14

**Table 8.18 Summary of comparison of experts' rules and JSM hypotheses.**

<u>Item</u>	<u>TB</u> <i>8 experts' rules.</i>	<u>CBr</u> <i>3 experts' rules.</i>	<u>HP</u> <i>No experts' rules.</i>	<u>Total</u> <i>11 experts' rules.</i>
Rules Confirmed	2	0	-	2
Rules Not Accounted	6	3	-	9
New Rules	3	0	1	4

Table 8.19 is a summary of the comparison of the experts' rules to the rules generated by both methods. The hypotheses of a context are a subset of the implications always. Hence this table is identical to Table 8.17 which summarises FCA implications. FCA implications are the superset of the two. The comparison was done to observe the differences between the two methods. In the conclusion it is suggested that hypotheses were found to be conservative.

**Table 8.19 Summary of comparison of FCA generated rules and experts' rules.**

<b><u>Item</u></b>	<b><u>TB</u></b> <i>8 experts' rules.</i>	<b><u>CBr</u></b> <i>3 experts' rules.</i>	<b><u>HP</u></b> <i>No experts' rules.</i>	<b><u>Total</u></b> <i>11 experts' rules.</i>
Rules Confirmed	8	1	-	9
Rules Not Accounted	0	2	-	2
New Rules	8	2	4	14

**Table 8.20 CDSS performance, for all illnesses.**

<b><u>Study Illness</u></b>	<b><u>Performance on Test data</u></b> <i>(original experts' rules)</i>	<b><u>Performance on Test data</u></b> <i>(new rules)</i>
Tuberculosis	80%	90%
Chronic Bronchitis	-	-
Hypertension	-	70%

Note: Hypertension did not have expert rules when data collection happened, so there are no performance values for original experts' rules. The performance for bronchitis was not done because its dataset was too small to split into two.

## **8.6 Chapter Conclusion**

The conclusion of this chapter is presented in two parts. The first is a general discussion of the results, and then an account is given of the discussions held with the experts on the results.

### **8.6.1 Discussion of Results**

In the beginning the motivation to use FCA and experts' knowledge to build the CDSS was that FCA data analysis may perhaps discover some things that the experts may have overlooked. Or

at least the experts can review their expertise against actual field data which has been analysed by FCA. A system like the CDSS cannot be built using FCA data analysis techniques only, involvement of experts is very important.

From Table 8.19, it was found that the FCA methods subsumed 9 out of 11 original experts' rules. This could be interpreted as 9 out of 11 experts' rules were accounted for by rules generated by FCA methods. The number of experts' rules not accounted for is 2. This means that the rules generated by FCA methods subsumed most of the rules that were originally proposed by the experts.

FCA methods would have been considered a failure had they contradicted most of what the experts had suggested, or if they had failed to generate some of the rules originally proposed by the experts. This is because it is expected that most of the rules suggested by the experts are part of accepted and valid medical knowledge. FCA rules then should agree in some cases, and in others they may result in refinement.

14 new rules were also generated by the FCA methods, Table 8.19. This is 127% more than the 11 original experts' rules. Of course there were no original experts' rules for hypertension, so perhaps the figure would be lower if there were original experts' rules for hypertension. For tuberculosis the number of rules generated by the FCA methods was (coincidentally) equal to the number of rules originally suggested by the experts: 8. For bronchitis it was 2.

Hypotheses are a subset of the implications of a context. The comparison of the two methods was done because one of the methods may perhaps yield better results. FCA implications subsumed (accounted for) more experts' rules for all the illnesses and had a low number of rules not accounted for, compare Table 8.17 and Table 8.18. On the other hand JSM hypotheses subsumed fewer rules and had a high number of experts' rules that were not accounted for. The JSM method is known to make fewer predictions than FCA implication. The terminology that has been used in the study to compare FCA implications and JSM hypotheses is that JSM hypotheses are "conservative" or that FCA implications are "opportunistic". (Kuznetznov and Ganter (2000) used the term "courageous" in the latter context. However, that seems to ascribe a positive quality to the rules which may not always be warranted, as will be explained below.)

By conservative it is suggested that JSM hypotheses require more evidence than implications and that as a result they tend to make fewer predictions. In the case of the CDSS, the evidence

referred to is the number of symptoms required in order to make a positive prediction. For example, in Table 8.5 of tuberculosis, new rule number 12 has only one symptom in its premise, {BS}, and has a corresponding implication. But there is no corresponding JSM hypothesis for this rule. There is actually no JSM hypothesis of the tuberculosis data which has one symptom only. This is partly because of the restriction that hypotheses have to be intents which are only found in positive examples. As a result they tend to have larger premises (more conjuncts), and there tends to be fewer of them. For example in the tuberculosis results, there are 8 non-redundant implications which conclude about tuberculosis. But there are only 3 JSM hypotheses which do the same. As a consequence the implications were able to subsume more experts' rules. It does not necessarily mean that the predictions made by implications will be correct always. This is perhaps because some of them have low support. When implications with low support are used (in patient assessment) it can result in many false positive predictions (and hence preference for characterizing such rules as opportunistic, rather than courageous). They are opportunistic in that they make many predictions based on small evidence, hence they may make many false positive predictions. But by doing so, they also make many correct predictions, as if to compensate. But being conservative also has disadvantages because it could imply that some positive predictions which should have been made are not made. It may remain undecided because it requires a large number of symptoms. This is an error too.

With respect to the problem of medical diagnosis by a program like the CDSS it is perhaps safer to suggest an illness when there is a worrying symptom even if later nothing is found, than to suggest that the patient is fine when in fact they are not. The JSM method by not predicting when it should could be interpreted as predicting no disease when in fact there was one. This discussion goes back to the ethical challenge of a health decision support system like the CDSS. What happens when it does not predict illness when there is one? So perhaps an opportunistic strategy than a conservative one should be used. It also needs to be checked that there can be no serious harm resulting on the patient if it fails to predict when it should have done so. There is greater risk of harm when a false negative prediction is made, than when a false positive prediction is made. But there is no strategy which guarantees that a problem will be found always when that it is the case. This applies to both the opportunistic strategy and the conservative strategy (Miller & Geissbulher, 1999). For example if it fails to predict tuberculosis when it was the case, there should be some backup mechanism. Something like, if you do not get better in five days go to the doctor. But at this stage it refers all patients to the clinic regardless of its diagnosis (conclusion).

This "conservative" tendency of JSM hypotheses is also associated with the fact that support for JSM hypotheses is mostly high, while for some implications it is low. For example, in the tuberculosis results the following was found. The implication  $BS \implies TB$ , has support 1. While support for the three JSM hypotheses for the same dataset is 10 for two of them and 11 for the other, see Table 8.5.

On the whole the FCA methods (both implications and JSM hypotheses) generated more opportunistic rules than did experts. That is, the rules of the experts used more symptoms while those generated by the FCA methods used less symptoms. For example in Section 8.2.1 for tuberculosis, experts' rule number 3 requires the following symptoms {PC WL NS} while the corresponding FCA methods' rule, rule number 14, Table 8.5 requires only {WL NS}. There is an example though where an implication has been more conservative than an experts' rule. New rule 4 of bronchitis Table 8.11.

The performance of the experts' rules was 80% for tuberculosis. Table 8.20. When the new rules were added the performance went up to 90%. The performance of the new rules for hypertension was 70%. There were no original rules for hypertension, hence there was no value for performance for original experts' rules. There was no test dataset for bronchitis due to limited data, hence no performance values for bronchitis. But it can be said that the performance of the CDSS went up after addition of the new rules based on the results for tuberculosis. The general performance was above 70% for all the illnesses, for the experts' rules as well as the new FCA methods' rules. This matched the target which had been set of 70% as a reasonable performance value.

Perhaps it can be said that the new rules are promising. It can also be said that the CDSS can help many people if this were to remain its performance. But it must still be determined what adverse effects can arise from failure to diagnose when it should.

Another interesting observation on the performance of the CDSS is the following. Its predictions tended to fail by not making a positive prediction when they should have. This is the case for both the experts' rules as well as the FCA methods' rules, see Table 8.7 and Table 8.16 of CDSS performance for tuberculosis and bronchitis respectively. Most of the CDSS positive predictions were correct though. This observation could perhaps steer the experts to consider deriving more opportunistic rules than currently is the case. Perhaps the data analyst could also configure the FCA methods to generate more opportunistic rules.

## 8.6.2 Experts' Discussion

This section is about the discussion of results of data analysis by the experts. Firstly, the chest conditions are considered (tuberculosis and bronchitis), and then hypertension.

### 8.6.2.1 Tuberculosis and Bronchitis

The experts noted that tuberculosis was a very serious problem and diagnosis using even fewer symptoms, what has been described as opportunistic rules, is acceptable. FCA was found to be valuable because it raised a discussion on how best to approach the problem of diagnosis by the CDSS. However, as noted in the CDSS diagnosis strategy Chapter 2 Section 2.3.3, the doctor still has to work through a differential diagnosis list when the patient arrives with a problem. It only suggests to the patient to seek help as they may be at risk of serious illness, for example tuberculosis. This could be interpreted to mean that tuberculosis would be (high up) on the doctor's differential diagnosis list.

The experts noted that even the laboratory tests that are done to test for tuberculosis are not 100% accurate. Sometimes a doctor would continue to treat a patient even if some of the tests (like the sputum test) are negative. This is a factor to be considered when selecting a gold standard which is used for measuring performance. The degree to which this affects the study is unknown, but it was felt that the usefulness of the CDSS or the conclusions that are drawn by the study would not be greatly affected.

The experts noted that the performance of the CDSS needs to be checked even for the cases where tuberculosis was not found but where some of the serious tuberculosis symptoms were present, for example persistent cough. Some other serious illness may have been found. The experts felt that even if the rules had suggested tuberculosis, but something else which was also serious were then found, this should be considered a success too. This is because its aim is not so much accurate diagnosis but to get the patient to seek help if there is a problem.

The experts also noted that some of the new rules are derived from implications with low support. For example, new rule 12, Table 8.5 of tuberculosis, is given by  $BS \rightarrow TB$ . New rule 12 is only supported by one (1) patient out of 21 that participated. But it was observed that BS (Bloody Sputum) is a very serious symptom.

The rules for bronchitis were found to be acceptable. The general discussion above is also applicable.

### **8.6.2.2 Hypertension**

There were no original experts' rules for hypertension but it was agreed by the experts that the combinations suggested by FCA methods were acceptable. It was noted by the experts that quite often hypertensive patients do not experience any obvious symptoms. Indeed, sometimes hypertension is found when a patient is being treated for something else. As a result, it is sometimes called the silent killer. The only way to find out if someone has hypertension is to do a blood pressure check. There are mild symptoms as well as risk factors that predispose someone to hypertension. The value of the CDSS, it was noted, was to identify and list explicitly some of these minimum symptoms and risk factors.

## Chapter 9

### Conclusion

#### 9.1 Summary of Study

The study was about evaluation of the knowledge base of a consumer health expert system called the CDSS program using FCA methods. CDSS means: Consumer Decision Support System (CDSS) for Important Common Ailments in South Africa. Its knowledge base consists of production rules which were provided by medical experts. The rules are of the relationship between symptoms and illnesses. This is in order to help members of the public to assess their own health in order to know when to seek professional medical help.

A field prototype of the CDSS program was developed and completed for use in the study. The experts on the development team included a chest specialist and a gastroenterologist who are both members of faculty at Medunsa Medical campus of University of Limpopo in South Africa. There was also a nursing sister with more than thirty years of experience including primary health care.

The CDSS program operates by prompting the patient (consumer) to enter his or her symptoms. Thereafter it assesses the potential that the user may be suffering from one of the common ailments on its list, and then recommends an action. It uses the knowledge base in order to conclude. Its conclusion can be loosely considered to be a preliminary diagnosis. In the project it is loosely called a diagnosis. The use of diagnosis is not in the strict medical sense which has strong legal and ethical implications.

However, the CDSS only includes rules which make *positive predictions*. Rules which make *negative predictions* are excluded. For example, the following rule is not allowed. “If the user has a persistent cough of only three days, and no other symptoms or risk factors then no TB is present.” If negative conclusions are possible, then self treatment at home might be an option.

However, in the absence of negative predictions, the CDSS in its present form *always* recommends that the user should undergo further medical examination, either to confirm the suspected diagnosis at which the CDSS system has arrived, or to establish a diagnosis in the

absence of a CDSS-derived one. Even in this preliminary form, it is considered that the CDSS would still be useful. This is because it is assumed that most users of the CDSS do indeed have a problem – they have simply not yet decided to seek help. In practice, this tardiness in seeking help may be due to distance from medical facilities, lack of motivation, the all-to-human tendency to procrastinate, etc. Its role would be to give added encouragement to seek medical advice; to provide information about a likely diagnosis where possible; perhaps to direct a user to a medical establishment that has specialized facilities for certain illnesses when necessary, etc.

Later it can perhaps be further developed to include positive and/or negative rules that predict minor illnesses (e.g. a sore throat), for which users may self-treat at home without the need to go to the clinic. Clearly, however, such bypassing of medical opinion could be highly contentious and raise significant ethical questions. At this stage, therefore, there is no question of using it to usurp the role and responsibility of the medical profession. Whether it can ever be refined up to a point where society would sanction such a role for a non-human diagnostic system is a matter of speculation that is not within the scope of the present study. Eysenbach (2000) suggests that decision support systems for patients can be useful, but also states that it is not known exactly how consumers use this medical knowledge. This is a challenge for the project.

It is intended to be deployed in a location where members of the public can have easy access. The location could be a clinic or any other place where members of the public can have easy access to.

## **9.2 Summary of Study Design**

The aim of the study was to test the rules of the knowledge base of the CDSS against a sample of actual patient data using FCA.

The field prototype of the CDSS program was deployed in a primary health care (PHC) clinic where it was used to collect data on patients' symptoms. A PHC clinic was chosen as the data collection site as it would be easy to find patients with the symptoms and illnesses of the study. Study patients were chosen from among patients who had come for a visit to the clinic. Patient data was collected as follows. The CDSS was used to collect the patients' symptoms using the CDSS questionnaire. It was possible to use it to collect patient symptoms because it used the same questionnaire that would be used by the researcher. The CDSS illnesses questionnaire can be found in Appendix II. Thereafter the patient went to see the doctor. Then the doctor completed

the patient's data by providing the diagnosis. The data for each illness was formatted into a table which contained the symptoms for each patient as well as the treating doctor's diagnosis.

The actual data collection deviated from the above plan. This was done in order to be able to collect the required sample size within the time available for data collection. Positive examples were collected as follows. Hypertension patients were selected from known chronic cases who had come to the clinic for medication and scheduled visits. Tuberculosis patients were selected from among previously diagnosed patients who had come for medication. Treatment for tuberculosis takes about six months. Chronic bronchitis patients who were also selected from previously diagnosed patients. Then the interview was conducted by asking what the patient felt when initially ill. At the time of the interview the patients' condition may have improved due to medication. So it was relied on the patients' memory of their earlier symptoms. Negative examples were selected with the help of the doctors who identified those patients that had some symptoms of the illnesses but who were not diagnosed with the illnesses. For example for tuberculosis, it may be patients with a persistent cough but who are not found to have tuberculosis.

Then FCA methods were used to generate rules from the patient data. The rules generated by the FCA methods were then compared to rules provided by the experts as follows.

- *Rules Subsumed.* How many experts' rules have been subsumed by the FCA rules? These can be considered to be experts' rules which are accounted for by the FCA generated rules.
- *New Rules.* These are rules generated by the FCA methods which have no corresponding experts' rules. These rules may be considered enrichment to the CDSS if the experts accept them as valid.
- *Rules not Accounted.* These are experts' rules which have no corresponding FCA rules as defined above in the first bullet. These rules can be considered misses (omissions) by the FCA methods.

In addition, the new rules generated by the FCA methods were tested by measuring the performance of the CDSS at giving advice. By performance the following is meant. How many times did the diagnosis of the treating doctor coincide with that of the CDSS. The performance was measured twice. First it was measured with the original experts' rules only. Then the new rules generated by FCA methods were included, and the performance was measured again. The aim was to check whether the performance of the CDSS improved or not, as a result of inclusion of the new FCA methods generated rules. To do this the data was split into two sets. There was

a *training set* which was used for generating the FCA method rules. And there was a *test dataset* which was used to test its performance with the new rules included. The results of data analysis were then taken back to the experts for consideration in the iterative manner of knowledge engineering of expert systems. The experts had to make the decision as to what should be accepted or not from the FCA analysis.

In the treatment of tuberculosis the gold standard is a laboratory test. In the study the conclusion of the treating doctor was used however. This was accepted by the development team which included medical experts. It was however noted that it was a weakened standard. This also applied to chronic bronchitis. The treating doctor though was considered a very good standard for hypertension.

### **9.3 Study Proposition**

Preliminarily it was found that FCA methods can be used in the evaluation and development of the CDSS health expert system. The following method for evaluation of rules of a knowledge base of an expert system was proposed.

Rules are collected from experts and are implemented in the knowledge base of the expert system being developed. Field data is then collected and then FCA implications and the FCA formulation of the JSM method are used to generate rules from the data. The rules generated by using FCA methods are compared to the rules originally given by the experts. The results can be taken back to the experts for consideration in the iterative manner of knowledge engineering of expert system development.

The results of this study are described as preliminary because the data sample which was collected was limited in size. Next a summary of data analysis is made.

### **9.4 Summary of Data Analysis**

In Table 9.1, it is shown that the FCA generated rules have subsumed most of the rules originally proposed by the experts. Only two experts' rules were not subsumed by the FCA rules. In addition 14 new rules were generated by the FCA methods. This number is greater than the number of rules originally proposed by the experts. All the new rules were accepted by the experts.

**Table 9.1 Summary of comparison of FCA generated rules and experts' rules.**

<u>Item</u>	<u>TB</u> <i>8 expert rules.</i>	<u>CBr</u> <i>3 expert rules.</i>	<u>HP</u> <i>No expert rules.</i>	<u>Total</u> <i>11 expert rules.</i>
Rules Confirmed	8	1	-	9
Rules Not Accounted	0	2	-	2
New Rules	8	2	4	14

The performance of the CDSS after inclusion of the new rules compared favourably with the performance for the original experts' rules, Table 9.2. For tuberculosis the performance went up. Overall its performance was higher than the target of 70% which had been set. From this limited study, it can be said that preliminarily FCA has broadened the knowledge base and made it more sensitive.

**Table 9.2 Summary of CDSS performance.**

<u>Study Illness</u>	<u>Performance on Test data</u> <i>(original expert rules)</i>	<u>Performance on Test data</u> <i>(new rules)</i>
Tuberculosis	80%	90%
Chronic Bronchitis	-	-
Hypertension	-	70%

The comparison between FCA implications and the JSM hypotheses found that implications generated more new rules and confirmed more experts' rules, Table 9.1. From this, can it be asked whether implications are better than the JSM hypotheses? From the limited data, it can be said that FCA implications were more valuable.

A concern that may be raised is that the support for implications varies widely. Sometimes it is high, and sometimes it is low. Then the following question can be asked: is it good to use

implications which have low support. For example New Rule 12 of the tuberculosis dataset, Table 8.5 is given by:  $BS \rightarrow TB$ . New Rule 12 is only supported by one patient out of 21 that participated. It corresponds to an implication, but has no corresponding JSM hypothesis. It was found that the JSM hypotheses were generally of high support, for all the illnesses.

In the literature (Blinova et al, 2003) it is suggested that if the definition of a hypothesis is ‘weakened’, then more hypotheses may be generated from data. This could lead to better performance. For example, if a positive hypothesis is defined as an intent in which 70% of the cases covered exhibit the condition, then this would be the case. This is called the *confidence*, and it is different to *support*. In words this can be interpreted loosely as follows. That which is used to determine the target attribute can also be present in negative examples. But it must appear in more positive examples than it appears in negative examples (or at least it must appear in an agreed minimum number of examples). The rigid definition of a (positive) hypothesis states that a (positive) hypothesis is an intent which appears only in positive examples. That is, its confidence is rigidly set at 100% confidence.

In the case of hypertension there had been no initial experts’ rules. But using the patient data that had been collected it was possible to use FCA methods to generate rules which were accepted by the experts. This suggests another use for FCA in the development of rules of an expert system. If experts were to provide a dataset, in some instances then FCA could be used to generate rules where none existed.

Another interesting observation on the performance of the CDSS is the following. Its predictions tended to fail by not making a positive prediction when they should. This is a weakness, though the rate of failure was not very high. This is the case for both the experts’ rules as well as the FCA methods’ rules. This observation could perhaps steer the experts to consider deriving more opportunistic rules than currently is the case. And perhaps for the data analyst to tune the FCA methods to generate more opportunistic rules.

## 9.5 Discussion: Consumer Health Informatics

In the beginning of the project after an informal survey it was found that there are few CHI applications developed for problems specific to South Africa. It was also found that CHI was a relatively new field in South Africa, and that usage of CHI applications for the general population was still low. Around the world CHI has been recognized as having potential. As a

result there are many research endeavors in the field of CHI and the field is predicted to grow. A good example is the NHS Direct Online advice service in the United Kingdom which has been operating online since 1999. There is also another ambitious project in the United Kingdom called NHS Connecting for Health which is planned, it is a national health informatics project with a multi-billion pound budget. NHS Connecting for Health has a substantial CHI part. Based on this it was also felt that CHI can also play a role in health service delivery in South Africa. The relatively high performance of the CDSS can be viewed as preliminary support for a system like the CDSS in South Africa.

It was designed to meet information needs of members of the public for some common but important (serious) illnesses found in South Africa. The program can be deployed in information stations intended for the public that the government plans to make available. Some examples are the multi-purpose community computer centres (MPCC), or the Public Internet Terminals (PIT) of the Department of Communication (Post Office).

The following two challenges to widespread usage of the CDSS program were identified. Firstly the levels of education and sophistication of the target users are low. Secondly the degree of access to information communication technology (ICT) facilities is very low in South Africa. Most people who have access are relatively well educated and well-off. But it is expected that ICT will be used increasingly in people's lives as is the case in developed countries like England and the USA. It is envisaged that access will increase for the general population, and the levels of computer literacy will also go up.

The CDSS software program is a comparatively small endeavor which did not demand substantial resources to build. It was built by the knowledge engineer, as a student-related project, with input from the other team members at intervals. Perhaps it could also be an example of how to build a software program that is relevant to South Africa. South Africa is a developing country, looking for viable solutions to some of her most pressing problems. The solutions should be developed using limited resources while making a difference.

## 9.6 Discussion: Expert Systems

Expert systems were used widely in Artificial Intelligence programs in the nineteen seventies and nineteen eighties, but since then their profile has come down. Nevertheless in the CDSS study it was found that they are still a very useful development style. There are many systems which were built in the expert system style back then. For example MYCIN which was used as a case study. MYCIN is a very well documented research project. MYCIN's architecture is well documented, as well as many evaluation studies that were done on it. Good documentation is not always easy to find with many technologies. Expert systems are still used today. For example, the insurance industry in South Africa routinely relies on expert system technology to do a first level screening of client applications, filtering out the simple cases, and leaving it to consultants to deal with non-standard applications.

Rule-based systems are also still used, and some examples were found. The CDSS program is a rule-based expert system. The following benefit of rules was identified. Rules can facilitate knowledge acquisition; this is because their syntax and semantics can be understood both by the engineer and the experts. Normally there is a wide gap between the description of knowledge in the domain, and the translation into computer code. This aspect of rules was found to facilitate development of the system and to improve communication.

## 9.7 Gaps

Two major studies are to follow the current study. The first is to evaluate aspects related to the human computer interaction (HCI) of the CDSS program. An aspect to consider is the translation of medical domain knowledge, which is technical, for use by the lay man. It was said that the target audience for the CDSS program is of low education and sophistication. It needs to be checked how this can be overcome.

The second major study would evaluate its social impact. This study would measure its potential to make a difference in the health of the community.

The following gap was identified in its advice. The user is referred to a clinic always even if they may not have serious symptoms. It could perhaps be developed to include rules for some minor

ailments which do not need a visit to a health provider, in that case the user may be able to do self care at home.

With regards to gaps relating to the FCA part of the study, it needs to found out the effect of FCA on a larger more widespread sample than the one employed in the current study. The aim of this would be twofold. Firstly the soundness of the methodology should be further investigated. Secondly it should be determined whether further refinement can be achieved, and whether more illnesses can be included.

Another FCA method may perhaps be considered, for example *association rules* (Zaki & Ogihara, n.d.). Association rules are a more statistical variation of the two methods employed in the current study. Perhaps finer rules may be found as a result. In the study association rules were not used because they required a statistically significant dataset than was possible to collect in the study.

## Bibliography

- Aikins, J. S. (1983). Prototypical Knowledge for Experts Systems. *Artificial Intelligence*, 20(2): 163-210.
- Avidon, V. V. & Pomerantsov, A. B. (1982). Structure-activity relationship oriented languages for chemical structure representation. *J. Chem. Inf. Comput. Sc.* 22: 207-214.
- Berners, E. S. (Editor). (1999a). *Clinical Decision Support Systems: Theory and Practice*. New York: Springer-Verlag.
- Berners, E. S. (1999b). Testing System Accuracy. In *Clinical Decision Support Systems: Theory and Practice*. Edited by Berners, E.S. New York: Springer-Verlag.
- Blinova, V. G. Dobrynin, D. A. Finn, V. K. Kuznetsov, S. O. & Pankratova, E. S. (2003). Toxicology Analysis by Means of the JSM-method. *Bioinformatics*, 19(10): 1201-1207.
- Bouhaddou, O. Lambert, J. & Miller, S. (1998). Consumer Health Informatics: Knowledge Engineering and Evaluation Studies of Medical HouseCall. In *proceedings of the American Medical Informatics Association Symposium 1998*. 612-616.
- Bradshaw, D. Nannan, N. Laubscher, R. Groenewald, P. Jourbert, J. Nojilana, B. Norman, R. Pieterse, D. & Schneider, M. (2000). *Estimates of Provincial Mortality, Summary Report*. South African Medical Research Council, Burden of Disease Unit. Available from: <http://www.mrc.ac.za/bod/reports.htm> . (Accessed April 2006).
- Buchanan, B. G. & Shortliffe, E. H. (Editors). (1984a). *Rule-based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project*. Reading, Massachusetts: Addison-Wesley.
- Buchanan, B. G. & Shortliffe, E. H. (1984b). A Model of Inexact Reasoning in Medicine. In *Rule-based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project*. Edited by Buchanan, B. G. & Shortliffe, E. H. Reading, Massachusetts: Addison-Wesley.

Buchanan, B. G. & Shortliffe, E. H. (1984c). The Problem of Evaluation. In *Rule-based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project*. Edited by Buchanan, B. G. & Shortliffe, E. H. Reading, Massachusetts: Addison-Wesley.

Buschmann, F. Meunier, R. Rohnert, H. Sommerland, P. & Stal, M. (2000). *Pattern-Oriented Software Architectur, a System of Patterns*. Chichester, New York: John Wiley and Sons.

Carty, B. & Kenney, K. (2006). Consumer Health Informatics in Primary Health Care. In *Consumer-Centred Computer-Supported Care for Healthy People*. Edited by Park, H. Murray, P. & Delaney, C. Conference proceedings of the 9<sup>th</sup> International Congress on Nursing Informatics held in Seoul South Korea in June 2006. Conducted by the Korean Society of Medical Informatics. Amsterdam: IOS Press.

*CLIPS, rule-based programming language*. (2006). (Computer software). Available from: <http://www.ghg.net/clips/clips.html> . (Accessed January 2006).

*Concept Explorer*. (2006). (Computer software). Available from: <http://sourceforge.net/projects/concexp>. (Accessed April 2006).

Department of Health, Republic of South Africa. (2003). *Standard Treatment Guidelines and Essential Drugs List*. Pretoria: Department of Health.

Eminovic, N. Wyatt, J. C. Tarpey, A. M. Murray, G. & Ingrams, G. J. (2004). First Evaluation of the NHS Direct Online clinical Enquiry Service: A Nurse-led Web Chat Triage Service for the Public. *Journal of Medical Internet Research*. 6(2):e17. Available from: <http://www.jmir.org/2004/2/e17/> . (Accessed October 2005 ).

Eysenbach, G. (2000). Recent Advances, Consumer Health Informatics. *British Medical Journal*, 320: 1713-1716.

Eysenbach, G. Powell, J. Kuss, O. & Sa, E. R. (2002). Empirical Studies Assessing the Quality of Health Information for Consumers on the World Wide Web: A Systematic Review. *Journal of the American Medical Association*, 287(20): 2691-2700.

Finn, V. K. (1983). On Machine-Oriented Formalisation of Plausible Reasoning in the Style of F. Backenn – J. S. Mill. *Semiotika i Informatics*, 20: 35-101.

Ganter, B. & Wille, R. (1999). *Formal Concept Analysis: Mathematical Foundation*. Berlin: Springer-Verlag.

Ganter, B. & Kuznetsov, S. (2000). Formalising Hypothesis with Concepts. In Conference *Proceedings of the 8<sup>th</sup> International Conference on Conceptual Structures*. Edited by Ganter, B. & Mineau, G. 1867: 342-356.

Garlan, D. Allen, R. & Ockerbloom, J. (1994). Exploiting Style in Architectural Design Environments. In Conference proceedings of the ACM SIGSOFT 1994, Symposium on Foundations of Software Engineering. Available from:  
<http://citeseer.ist.psu.edu/cache/papers/cs/4760/http:zSzzSzalmond.srv.cs.cmu.edu:zSzafszSzczszSzprojectzSzablezSzftpzSzaesop-fse2.pdf/garlan94exploiting.pdf> . (Accessed January 2006).

Goodman, K. W. (1999). Ethical and Legal Issues in Decision Support. In *Clinical Decision Support Systems: Theory and Practice*. Edited by Berners, E.S. New York: Springer-Verlag.

Guigues, J. L. & Duquenne, V. (1986). Familles minimales d'implications informatives resultant d'un tableau de donnees binaires. *Math. Sci. hum.*, 24(95):5-18.

Harmon, P. & King, D. (1987). *Artificial Intelligence in Business: Expert Systems*. New York: John Wiley and Sons Inc.

Health24 consumer health website in South Africa. Available from: <http://www.health24.com/>. (Accessed August 2006).

Horner, V. Hanmer, L. & Mbananga, N. (2006). A Consumer Decision Support System for Common Health Ailments in South Africa. In *Consumer-Centred Computer-Supported Care for Healthy People*. Edited by Park, H. Murray, P. & Delaney, C. Conference proceedings of the 9<sup>th</sup> International Congress on Nursing Informatics held in Seoul South Korea in June 2006. Conducted by the Korean Society of Medical Informatics. Amsterdam: IOS Press.

Park, H. Murray, P. & Delaney, C. (Editors). (2006). *Consumer-Centred Computer-Supported Care for Healthy People*. Conference proceedings of the 9<sup>th</sup> International Congress on Nursing Informatics held in Seoul South Korea in June 2006. Conducted by the Korean Society of Medical Informatics. Amsterdam: IOS Press.

Jackson, P. (1999). *Introduction to Expert Systems*. Harlow, England: Addison-Wesley.

*JESS, a Java-based rule programming environment*. (2006). (Software programme). Available from: <http://herzberg.ca.sandia.gov/jess/>. (Accessed October 2006).

Jimison, H. B. & Sher, P. P. (1999). Decision Support for Patients. In *Clinical Decision Support Systems: Theory and Practice*. Edited by Berners, E.S. New York: Springer-Verlag.

MedlinePlus, health consumer medical encyclopedia. Available from: <http://www.nlm.nih.gov/medlineplus/encyclopedia.html/>. (Accessed June 2006).

Miller, R. A. & Geissbuhler, A. (1999). Clinical Diagnostic Decision Support Systems: An Overview. In *Clinical Decision Support Systems: Theory and Practice*. Edited by Berners, E.S. New York: Springer-Verlag.

Newell A; (1984); (Forward to) *Rule-based Expert Systems: the MYCIN Experiments*, Buchanan B & Shortliffe E (Eds); Addison-Wesley Reading, Massachusetts.

NHS Connecting for Health Fact Sheet. Available from: [http://www.connectingforhealth.nhs.uk/publication/nhs\\_cfh\\_factsheet.pdf](http://www.connectingforhealth.nhs.uk/publication/nhs_cfh_factsheet.pdf) . (Accessed January 2006).

NHS Direct Online website. Available from: <http://www.nhsdirect.nhs.uk/> . (Accessed January 2006).

Pallesen, B. Engberg, A. & Balach, A. (2006). Developing e-Health Information by Empowerment Strategy. In *Consumer-Centred Computer-Supported Care for Healthy People*. Edited by Park, H. Murray, P. & Delaney, C. Conference proceedings of the 9<sup>th</sup> International Congress on Nursing Informatics held in Seoul South Korea in June 2006. Conducted by the Korean Society of Medical Informatics. Amsterdam: IOS Press.

Pople, H. E. (1982). Heuristic Methods for Imposing Structure on Ill-Structured Problems: The Structuring of Medical Diagnostics. In *Artificial Intelligence in Medicine*. Edited by Szolovits, P. Boulder, Colorado: Westview Press.

*Predictive Toxicology Challenge*. (2001). Available from: <http://www.predictive-toxicology.org/ptc/#DataEng>. (Accessed October 2006).

Project EAGLE, legal advice expert system. (n.d.). (Software program). Available from: <http://www.projecteagle.org.uk/>. (Accessed May 2006).

*QUDA, data analysis software*. (2006). (Software program). Available from: <http://kirk.intellektik.tu-darmstadt.de/~quda/QDA.zip>. (Accessed June 2006).

Rumbaugh, J. Blaha, M. Premerlani, W. Eddy, F. & Lorenzen, W. (1991). *Object-oriented Modeling and Design*. Englewood Cliffs: Prentice-hall.

Sailer, M. (2006). The Development of a Computerised Case-based System for Supporting and Measuring Clinical Decision-Making-Process Using the S-CDNT. Poster presentation 9<sup>th</sup> World Congress on Self-care Deficit Nursing Theory, 2006.

Sebesta, R. W. (1999). *Concepts of Programming Languages*. Reading, Massachusetts: Addison Wesley .

Shabo, A. (2006). A Global Socio-economic-medico-legal Model for Sustainability of Longitudinal Electronic Health Records. *Methods of Information in Medicine*. 45:240-245.

Shortliffe, E. H. (1984). Details of the Consultation System. In *Rule-based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project*. Edited by Buchanan, B. G. & Shortliffe, E. H. Reading, Massachusetts: Addison-Wesley.

Silberschatz, A. Henry, F. K. Sudarshan, S. (2002). *Database System Concepts*. McGraw-Hill.

Smith, R. (1997). The Future of Healthcare Systems. *British Medical Journal*, 314: 495-1496.

Schwartz, K. L. Roe, T. Northup, J. Meza, J. Seifeldin, R. & Neale, A. V. (2006). Family Medicine Patient's Use of the Internet for Health Information: A MetroNet Study. *Journal of the American Board of Family Medicine*, 19: 39-45.

Stumme, G. (2003). *Formal Concept Analysis, Methods and Applications in Computer Science*. Lecture Notes, Otto-von-Guericke-Universitat Magdeburg. Available from <http://...> (Accessed November 2005).

*Verification, Validation and Evaluation*. (n.d.). Notes. Available from: <http://www.tfhr.gov/advance/vve/vve4a.htm> . (Accessed January 2007).

van Bommel, J. H. & Musen, M.A. (Editors). (1997). *A Handbook of Medical Informatics*. Heidelberg, Germany: Springer.

Wikipedia, general encyclopedia. Available from <http://www.wikipedia.org/> . (Accessed August 2006).

Wong, T. K. S. (2006). Feeling the Digital Pulse: Consumer Centred Approach to Individual Health Profiling. In *Consumer-Centred Computer-Supported Care for Healthy People*. Edited by Park, H. Murray, P. & Delaney, C. Conference proceedings of the 9<sup>th</sup> International Congress on Nursing Informatics held in Seoul South Korea in June 2006. Conducted by the Korean Society of Medical Informatics. Amsterdam: IOS Press.

Wyatt, J. & Keen, J. (1998). The NHS's New Information Strategy. *British Medical Journal*, 317: 900.

Yu, V. L. Fagan, L. M. Bennett, S. W. Clancey, J. W. Scott, A. C. Hannigan, J. F. Blum, R. L. Buchanan, B. G. & Cohen, S. N. (1984); *An Evaluation of MYCIN's Advice*; In *Rule-based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project*. Edited by Buchanan, B. G. & Shortliffe, E. H. Reading, Massachusetts: Addison-Wesley.

Zaki, M. J. & Ogihara, M. (n.d.). *Theoretical Foundations of Association Rules*. Technical Report, University of Rochester, New York. Available from: <http://citeseer.ist.psu.edu/cache/papers/cs/191/ftp:zSzzSzftp.cs.rochester.edu/zSzpubzSzpaperszS>



[zsystemszSz98.DMKD.Theoretical foundations of association rules.pdf/zaki98theoretical.pdf](#) .

(Accessed June 2006).

## Appendix I: Glossary of Terms

### **Application Programming Interface (API)**

An API is a set of routines, protocols and tools used for building software applications. Most software environments provide API. For example Microsoft Windows operating system provides WinAPI in order to expose its functionality for use by software developers. This is so that software developers can develop programmes that can be used. Some examples of programs that can be developed using API are data entry programs, auditing software programs, or database access programs. Another example of an API is JESS, which is used in the development of the CDSS program. JESS is a set of routines that allow a Java program to access CLIPS which is a rule programming environment.

### **Bronchitis**

Chronic bronchitis is an inflammation of the bronchi. The bronchi are the main air passages in the lungs. Bronchitis is characterised by a productive cough which persists for a long period of time, or which occurs repeatedly.

### **CLIPS**

CLIPS is a programming language which is used for developing expert systems. Its basic programming construct are rules. It has the following features and characteristics: knowledge representation, portability, full documentation and is low cost (it is a free tool available over the Internet).

### **Concept Explorer**

Concept explorer is an automated data analysis tool used for generating implications. Its interface is GUI based, it has a user manual. It is a free tool that can be found at Source-Forge on the Internet at <http://sourceforge.net/projects/conexp>, its size is less than two megabytes. It is Java-based. In order to use it at least JRE version 5 is needed. It is possible to draw a line diagram of a context, which can be stored as a jpeg image for printing. It can also be used to generate association rules which are another FCA technique. Concept explorer supports comma separated value (CSV) type for data input.

## **Concept Lattice**

For the purposes of this dissertation the term formal concept lattice can be considered a synonym for FCA.

## **Consumer Health Informatics**

Consumer health informatics (CHI) is a subfield of health informatics which is interested in the development and research of information systems which meet the health needs of members of the public. According to the parlance of the domain of CHI, members of the public as opposed to health practitioners are called consumers. An example of a health practitioner is a medical doctor or nurse. An example of a consumer health informatics application is the Health24 website in South Africa. It can be found at <http://www.health24.com/>.

## **Diagnosis**

Diagnosis is a process of identifying a medical condition by its signs, symptoms and from results of various diagnostic tests and procedures. The conclusion reached through this process is also called a diagnosis.

## **Differential diagnoses**

Before a medical condition can be treated it must be identified. The physician begins by collecting the patient's medical history, and examining the patient, and inquiring on family history. Then the physician lists the most likely causes, which must be investigated. The list of potential illnesses that may be the cause are called the differential diagnoses.

## **Expert System**

An expert system is a programme which reasons with a vast store of knowledge derived from experts which is used for problem solving. It is developed using the methodology of knowledge engineering and tools of expert systems.

## **Field Prototype**

A prototype evolves through a number of stages. The first prototype is normally used by the development team only, and is usually very primitive. The first prototype may not be safe or appropriate for use by users. As the development process advances more sophisticated and better prototypes are built until there is one which can be tested with the users. The testing of the prototype with the users is normally done in a controlled environment. A field prototype is

intended for use by the users in the actual deployment environment, outside of a lab or other such tightly controlled environment.

### **Formal Concept Analysis (FCA)**

FCA is a field of applied mathematics whose central notions are a *formal context* and a *formal concept*. A formal concept is based on the philosophic intuition that a concept (or a thing of life), may be considered to have attributes as well as objects that are common to those attributes. For example in context of biological life there is the following. A valid concept in life could consist of the following attributes: {Need Oxygen, Need Water, Produce Carbon Dioxide}. The objects associated with these attributes are: {Animal Kingdom, Plant Kingdom}. The adjective formal is used to imply that the context and concept are used in a mathematical sense rather than in the broad everyday sense, though the inspiration is from the everyday sense.

### **Health Informatics**

Health informatics can be loosely defined as the field interested in development and research of information systems which support medical practice. Some examples of health informatics endeavours and applications are: electronic patient records, clinical decision support systems, medical image recognition software, developing standards for electronic means of health and medical data exchange.

### **Hypertension**

Hypertension also means high blood pressure. It occurs when the pressure against the walls of the blood vessels is consistently high. It may have some symptoms, but may also not present with symptoms. It is important that everyone has their blood pressure checked always. Some of the severe problems associated with hypertension are stroke, and kidney damage.

### **Implications**

FCA implications are used in the study and have been called an FCA based data analysis method. An implication has the following format  $A \rightarrow B$ . A and B are attribute sets of the context under consideration. For example in the context of patient data for tuberculosis the following was an implication

{Weight Loss, Night Sweat}  $\rightarrow$  {Tuberculosis}. It can be interpreted as follows. If a patient presents with weight loss and extreme night sweat, then tuberculosis can be suspected.

## **Information system**

Information system is a term which is used loosely to refer to a software programme. Another term for a software programme is *software application*. The field of software engineering is also often referred to as the field of information systems. In South Africa there is a distinction between the field of computer science and information systems. Computer science is geared more towards the mathematical parts of software technology, while information system is geared more towards the “softer parts” which involve more of process and less of the mathematical angle. In Europe another term *informatics* is used synonymously to computer science, hence the area of health informatics. These distinctions are not always very clear, but they do not prevent functioning either. The enquirer quickly adjusts to the various loose usages. Effort has been made to be consistent, so that the reader of the dissertation does not get lost.

## **Human computer interaction (HCI)**

Human computer interaction is the field of computer science which is interested in issues of user interface design. As might have been experienced by many computer users, some programs are very complex. Some others are meant to help special category users for example blind users. As a result HCI is a vast field.

## **Java**

Java is a general purpose programming language and environment created by James Gosling of Sun Microsystems. Java programmes require the JRE in order to run.

## **Java runtime environment (JRE)**

Informally the JRE can be described as the program needed in order to run Java applications. A program is normally written to run on a specific machine. For example a program can be written to run on Microsoft Windows, or to run on Linux. Normally a programme that is written for Windows will not run on Linux, and also the other way around. But a Java programme will run on any machine that has the JRE installed.

## **JSM Method**

The term JSM method is used in the project for short for: FCA formulation of the JSM-method. The JSM method is so named in honour of nineteenth century English philosopher John Stuart Mill who proposed schemes of inductive reasoning. John Stuart Mill suggested that common effects are likely due to common causes. The JSM method was originally formulated in the 1980s

in terms of predicate logic by Finn. The FCA formulation of this method is described by Kuznetsov in.

### **Knowledge base**

A knowledge base is a module of an expert system which stores the knowledge acquired from experts. The knowledge is stored in an encoded form. The encoding method may be rules or frames or any other knowledge representation method. The knowledge base is used for reasoning about a problem of the domain of the expert system.

### **Knowledge Engineering / Knowledge Engineer**

Knowledge engineering is a methodology used for development of expert systems. It involves a team in which the knowledge engineer and experts are central. The knowledge engineer works with the experts in order to extract knowledge to use in the expert system. The knowledge derived from the experts is codified in very specific ways, for example rule.

### **Malaria**

Malaria is a parasitic disease characterised by fever, chills and anemia. It is caused by parasites spread by malaria bearing mosquitoes.

### **Medical informatics**

This is a synonym for health informatics.

### **Medical Record**

A medical record is a systematic documentation of a patient's medical history and care. It allows health care providers to provide continuity of care to individual patients. It serves as a basis for planning patient care, documenting communication between the health care provider and other health professionals contributing to the patient's care, assisting in protecting the legal interest of the patient and health care providers' responsibility for the patient's care. In addition it can be used to collect data for research and education.

### **MYCIN**

MYCIN is a rule based expert system developed at Stanford University by the Stanford Heuristic Programming Project (SHPP) in the nineteen seventies.

## Peptic Ulcer

A peptic ulcer is an erosion of the lining of the stomach or duodenum. Small ulcers may cause no symptoms, while large ones may cause serious bleeding. Most ulcers occur in the first layer of the inner lining. A hole that goes all the way through is called a perforation, and is a medical emergency.

## Primary Health Care

In medicine primary health care refers to the first point of consultation for patients. It is normally based in the community, as opposed to a hospital. A synonym for primary health care is general practice.

## Production Rule

A production rule has a premise and a conclusion. The premise is a set of conditions under which the rule will be activated. When a rule is activated in an expert system, it may insert a new fact into memory or it may execute an action. An example of a rule in the CDSS programme is the following:  $\{ \textit{persistent occipital headache, dizziness} \} \rightarrow \{ \textit{hypertension} \}$ . In normal language it says that: if a person has a persistent occipital headache and experiences dizziness then suspect hypertension. The premise is on the left hand side, and the conclusion is on the right hand side in the notation of production rules. Rules originate from the intuition that humans solve problems by applying a set of rules of the form that has just been described.

## Prototype

When developing a product, the first example which is made is the prototype. It is an original model (example) of the product that exhibits the essential features of the later production model. It is used for purposes of testing by the development team.

## QUDA – Qualitative Data Analysis

QUDA is an automated data analysis tool used for generating JSM hypotheses. It has a GUI interface and a distribution of the software programme can be found at <http://kirk.intellektik.informatik.tu-darmstadt.de/~quda/QDA.zip>. QUDA like Concept Explorer requires the Java runtime to work. It supports a number of other data analysis techniques and data input techniques. The data input technique used in the project was the comma separated values (CSV).

## **Software Design Style / Software Architecture Style**

A software architectural design can be informally described as a gross decomposition of a system into its computational elements and their interaction. For this term software design and software architecture are used interchangeably. A software architectural style is an established software design, which has been used and found to be of value. Another way of looking at it is as an accepted design example.

## **Software Engineering**

Software engineering is a general area of computer science whose main aim is the development of software. According to the discipline of software engineering, software is developed systematically, with the development process happening according to predefined phases and plans. The main phases of the software engineering process are: *analysis, design and implementation, and maintenance.*

## **Study / Research Protocol**

The purpose of the research (study) protocol is to provide a clear and complete description of the purpose and benefits of the research. It describes the methodology involved, informed consent process, questionnaires and other materials to be used in the research. It describes any factors that are important to the research, like potential risks to study subjects etc.

## **Symptom / Sign**

A symptom is a sensation or change in health function experienced by a patient. It may be graded as strong, mild or weak. In medicine it is described as a subjective report as opposed to a sign which is objective evidence. An example of a symptom is a feeling of pain related by a patient, while an abnormal appearance of a retina observed by the doctor is a sign.

## **Tuberculosis**

Tuberculosis is a contagious bacterial infection caused by mycobacterium tuberculosis. The lungs are primarily involved, but infection can spread to other organs. Its main symptoms are a productive cough - sometimes of bloody sputum. Fever and night sweats are other symptoms.

## **User Interface**

A computer program is normally intended to be used by a human operator other than the software engineer that developed it. The point of access and interaction of the human operator with the computer is called the user interface. In human computer interaction the operator of the

programme is called the user of the programme. The user interface for computer programs today normally consists of dialog screens, which have inputs and options. Multimedia though is also used, for example voice prompts and animations, and even video.

### **User interface design**

This is a synonym for human computer interaction.

### **Wrapper Façade**

The wrapper façade pattern is used to encapsulate low level functions and data structures within higher level object-oriented class interfaces which are more robust, concise and maintainable. Here it is enough to say that a pattern is an example of a software design style.

## Appendix II: CDSS Questionnaire

### TUBERCULOSIS QUESTIONS

"Do you have a persistent cough, of more than two or three weeks?"

"When you cough, do you produce sputum?"

"Do you cough up a lot of sputum?"

"When you cough, what is the color of the sputum? Is it: BLOOD-STAINED, GREE/YELLOW, WHITE?"

"Have you lost weight since having these problems?"

"Do you sweat heavily at night since having these problems?"

"Do you have difficulty breathing, do you run out of breath easily?"

"Do you have a pain in the chest when coughing, or when breathing, or at any other time? Is the pain: SHARP, STABBING, DULL, NO-PAIN?"

"Do you have an appetite?"

"Have you been in regular contact with someone with tuberculosis?"

"Do you feel tired most of the time?"

### BRONCHITIS QUESTIONS

"Do you have a persistent cough, of more than two or three weeks?"

"When you cough, do you produce sputum?"

"Do you cough up a lot of sputum?"

"When you cough, what is the color of the sputum? Is it: BLOOD-STAINED, GREE/YELLOW, WHITE?"

"Have you smoked cigarettes for many years?"

"Do you feel tightness in the chest?"

"Do you have difficulty breathing, do you run out of breath easily?"

"Do you experience a wheezing chest, this is a sound from the chest that you can feel?"

#### PEPTIC ULCERS' QUESTIONS

"Do you have pain in the centre region of the stomach?"

"Does the pain irritate and gnaw almost like a bad toothache or worse?"

"Has this pain been bothering you for a long time?"

"Do anti-acids like rennies, make you feel better?"

"Do some foods make you feel better, like milk or any other foods?"

"Do some foods make you feel worse, like tomato or any other foods?"

"Do you experience water brash, this is a watery liquid from the stomach that comes to the mouth?"

"Do you get this pain when hungry, does it wake you up at night?"

"Do you pass some blood in the toilet, the sign of blood is stool that is dark and tarry?"

"Have you vomited blood, the sign of blood is dark coffee-colored vomit?"

"Have you been using any drugs for arthritis for a long time?"

"Have you regularly used any of the following drugs: ASPIRIN - dispirin or GRANDPA? "

"Do you smoke?"

"Do you drink alcohol?"

"Do you regularly eat spicy-hot or chilli foods?"

"Do you experience a lot of stress either at work or at home?"

#### HYPERTENSION QUESTIONS

"Do you experience a headache in the back of the head?"

"Has this headache been bothering you for a long time?"

"Does your headache occur often?"

"Do you often experience blurred vision?"

"Do you often experience dizziness?"

"Do you often experience nose bleeding?"

"Have you experienced any swelling of feet and ankles, and do you consider yourself overweight?"

"Does anyone in your family have high blood pressure?"

"Do you smoke?"

"Do you drink?"

"Do you experience a lot of stress either at work or at home?"

### MALARIA QUESTIONS

"Does this fever occur periodically, for example at certain times of the day?"

"Does the fever cause you to shake and perhaps shiver?"

"Does the fever leave you feeling tired?"

"Do you experience a low-grade fever?"

"Do you experience a gradual rise in temperature during the fever, and do you experience a sudden drop in temperature when the fever subsides?"

"Does the fever cause you to sweat?"

"Do you experience a headache?"

"Have you visited a malarial area recently? The following are some examples in Southern Africa: MPUMALANGA South Africa, SWAZILAND, MOZAMBIQUE, MALAWI."

"Have you been bitten by mosquitos during your visit to the malarial area?"

"Did you take anti-malarial drugs before travel?"

"Did you have a blood transfusion recently?"

## Appendix III: Experts' Rules

Experts' rules were acquired only for tuberculosis, chronic bronchitis and peptic ulcers. Data was not collected for peptic ulcers and malaria due to data collection limitations. The study went ahead with the three conditions (tuberculosis, bronchitis and hypertension) for which data had been collected, and it was found that the results of data analysis did meet the goals of the study.

### A. Tuberculosis

**Table 1 Key Tuberculosis (TB) Symptoms**

<u>Symptoms</u>
PERSISTENT COUGH ( <b>PC</b> )
SPUTUM PRODUCTION ( <b>SP</b> )
MUCO-PURULENT ( <b>MC</b> )
SPUTUM BLOODY ( <b>BS</b> )
CLEAR-SPUTUM ( <b>CS</b> )
WEIGHT LOSS ( <b>WL</b> )
EXTREME NIGHT SWEATS ( <b>NS</b> )
NO APPETITE ( <b>NA</b> )
CHEST PAIN ( <b>CP</b> )
SHORTNESS OF BREATH ( <b>SB</b> )
TUBERCULOSIS CONTACT ( <b>TC</b> )
TIREDNESS ( <b>TN</b> )

## Experts' Rules

1. PC SP BS WL → TB
2. PC SP BS NS → TB
3. PC WL NS → TB
4. PC SP BS TC → TB
5. PC SP BS CP NA → TB
6. PC SP BS SB → TB
7. PC WL CP SB → TB
8. PC SP BS CP TN → TB

## B. Chronic Bronchitis

**Table 2 Key of Bronchitis (CBr) Symptoms**

<u>Symptoms</u>
PERSISTENT COUGH ( <b>PC</b> )
SPUTUM PRODUCTION ( <b>SP</b> )
SPUTUM MUCO-PURULENT ( <b>MP</b> )
CHEST TIGHTNESS ( <b>CT</b> )
SHORTNESS OF BREATH ( <b>SB</b> )
WHEEZING CHEST ( <b>WC</b> )
SMOKING ( <b>SM</b> )

## Experts' Rules

1. PC SM SP CT SB → CBr
2. PC SM SP WC → CBr
3. PC SM SP MC → CBr

## C. Peptic Ulcers

**Table 3 Key of Peptic Ulcer (PU) Symptoms**

<u>Symptoms</u>	<u>Comment</u>
EPIGASTRIC PAIN ( <b>EP</b> )	
CHARACTERISTIC PAIN ( <b>CH</b> )	
FOOD INTOLLERANCE ( <b>FI</b> )	
WATER BRASH ( <b>WB</b> )	
HUNGER PAIN ( <b>HR</b> )	
ANTI-ACID RELIEF ( <b>AR</b> )	
BLEEDING ( <b>BL</b> )	
WORSENING FACTORS ( <b>WF</b> )	<i>smoking, alcohol, stress, spicy food</i>
NSAID ( <b>ND</b> )	<i>Non-steroidal anti-inflammatory drugs</i>

### Experts' Rules

1.  $EP\ CH\ FI \rightarrow PU$
2.  $EP\ CH\ AR \rightarrow PU$
3.  $EP\ CH\ HR\ BL \rightarrow PU$
4.  $EP\ CH\ WB \rightarrow PU$
5.  $EP\ CH\ ND \rightarrow PU$
6.  $EP\ CH\ WF \rightarrow PU$