Expert system adjudication of hospital data in HIV disease management

By:

Asma Joseph

Supervised by:

Prof Sonia Berman

Dissertation presented in partial fulfilment of the requirements for the degree of

Master of Information Technology in the

Department of Computer Science,

University of Cape Town

January 2012

Plagiarism declaration

"I know the meaning of plagiarism and declare that all of the work in thesis, save for that which is properly acknowledged, is my own".

Signed: ____________________

Asma Joseph

Abstract

HIV[i] Disease Management Programs (DMP's) are comprehensive programs that are designed to manage the HIV infected patient's treatment in an integrated manner. When an HIV infected patient is hospitalized, the hospital diagnosis may not necessarily be related to their HIV disease (e.g. Insulin-Dependent Diabetes Mellitus). However, because HIV is a progressive immune deficiency disease, it is more probable that the hospital diagnosis is HIV-related. Adjudication of hospital data is critically important for a HIV DMP. This is because key interventions are needed to assess how the hospital diagnosis impacts the patient's HIV disease. The HIV DMP that was investigated in this thesis has a current process during in which a knowledge expert (pharmacist) manually adjudicates hospital data of HIV patients. The adjudication has 2 stages: Stage 1 (plan for intervention which includes various checks and decisions on preliminary action plan) and Stage 2 (actual intervention with the doctor and mutually agreeing on an action plan).

A *HIV-expert system* was developed for the electronic adjudication of hospital data of HIV patients. This expert system uses a pure SQL approach to storing production rules, implementing forward chaining inference and recommending specific actions. The electronic adjudication of hospital data is compared with manual adjudication. The electronic adjudication has the same 2 stages as the manual adjudication, but in the electronic system, the expert system derives the preliminary action plan. Stage 2 in both processes, remains unchanged, as this is the actual intervention with the doctor. For the evaluation of the *HIV-expert system,* a questionnaire was completed by users of the manual system, comparing it with the electronic system.

The main findings are that by implementing the *HIV-expert system* which electronically adjudicates hospital data in an HIV DMP, the time for stage 1 in the electronic system is 53% faster than when the manual system is used. . It was found that the adjudication of hospital data in the manual mode had a 92% accuracy compared with 100% accuracy in automated mode using the *HIV-expert system.* Based on these findings, it can be concluded that the *HIV-expert system does* improve clinical interventions and operational efficiency in an HIV DMP.

---

[i] See Glossary

Acknowledgements

Firstly many thanks to my colleagues in the IT department at work for their assistance as well as my colleagues in the Clinical department for also being very willing to assist and extremely patient especially during the experiment testing part of this project.

Special thanks go to Prof Sonia Berman who supervised this project and provided excellent guidance as I journeyed through the ups and downs of this project.

Lastly, the most thanks goes to my family for supporting me not just through this project, but the through the entire MIT course. It has not been easy but your unwavering support has been an inspiration.

Glossary

Acquired Immunodeficiency Syndrome (AIDS)

A disease of the body's immune system caused by the human immunodeficiency virus (HIV). AIDS is characterized by the death of CD4 cells (an important part of the body's immune system), which leaves the body vulnerable to life-threatening conditions, such as infections and cancers [The Body 2011].

AIDS defining conditions

Any of a list of illnesses that, when occurring in an HIV-infected person, leads to a diagnosis of AIDS, the most serious stage of HIV infection. AIDS is also diagnosed if an HIV-infected person has a CD4 count less than 200 cells/mm^3, whether or not that person has an AIDS-defining condition. The Centers for Disease Control and Prevention (CDC) published a list of AIDS-defining conditions in 1993. The 26 conditions include candidiasis, cytomegalovirus disease, Kaposi's sarcoma (KS), *Mycobacterium avium* complex, *Pneumocystis jiroveci* pneumonia, recurrent pneumonia, progressive multifocal leukoencephalopathy, pulmonary tuberculosis, invasive cervical cancer, and wasting syndrome [The Body 2011].

Aid for Aids (AfA)

An HIV disease management program in South Africa.

Antiretroviral therapy (ART)

The specific medication which an HIV infected patient starts at a specific stage of their disease and adherence/compliance is essential and continued lifelong.

ART is divided into specific classes: Nucleoside Reverse Transcriptase Inhibitors, Non-Nucleoside Reverse Transcriptase Inhibitors and Protease Inhibitors.

Cytomegalovirus (CMV)

A herpes virus that can cause infections, including pneumonia (infection of the lungs), gastroenteritis (infection of the gastrointestinal tract), encephalitis (inflammation of the brain), or retinitis (infection of the eye), in immunosuppressed people. Although CMV can infect most organs of the body, HIV-infected people are most susceptible to CMV retinitis [The Body 2011].

Disease Management Program (DMP)

Comprehensive programs that are designed to manage specific diseases in an integrated manner.

Human Immunodeficiency Virus (HIV)

The virus that causes Acquired Immunodeficiency Syndrome (AIDS). HIV is in the retrovirus family, and two types have been identified: HIV-1 and HIV-2. HIV-1 is responsible for most HIV infections throughout the world, whereas HIV-2 is found primarily in West Africa [The Body 2011].

Kaposi's sarcoma (KS)

A type of cancer caused by an overgrowth of blood vessels, which causes pink or purple spots or small bumps on the skin. The condition can also occur inside the body, especially inside the intestines, lymph nodes, and lungs. When inside the body, KS can be life threatening. In people infected with HIV, KS is considered an AIDS-defining condition. A virus called Kaposi's Sarcoma herpes virus (KSHV) or human herpes virus 8 (HHV-8) is associated with Kaposi's Sarcoma [The Body 2011].

Opportunistic infections (OI)

An illness caused by any one of various organisms that occur in people with weakened immune systems, including people with HIV/AIDS. OI's that are common in people with AIDS include: Kaposi's Sarcoma, Cytomegalovirus, *Pneumocystis jiroveci* pneumonia (PCP), cryptosporidiosis, histoplasmosis, toxoplasmosis, other infections (parasitic, viral, and fungal) and some types of cancers [The Body 2011]. Kaposi's Sarcoma and Cytomegalovirus which are the 2 conditions discussed in this thesis.

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction

1.1 The HIV epidemic

The HIV epidemic has a devastating global impact. Recent statistics for people living with HIV indicate that Sub-Saharan Africa continues to bear a disproportionate share of the global HIV burden. This is because an estimated 22.5 million people living with HIV resided in Sub-Saharan Africa in 2009 (representing 68% of the global HIV burden) and about 34% of all people living with HIV resided in the 10 countries of Southern Africa in 2009. With an estimated 5.6 million HIV-positive people, South Africa continues to have the world's largest HIV epidemic [UNAIDS 2010].

Globally, gains continue to be made in the response to the global HIV epidemic. New HIV infections are falling, fewer people are dying of AIDS-related causes and more people with HIV are living longer. Evidence of this is that there were 2.7 million new HIV infections in 2010 (down 21% from the peak of the global epidemic in 1997) and at the end of 2010 an estimated 34 million people were living with HIV worldwide (up 17% from 2001) [UNAIDS 2011].

The positive impact that providing HIV treatment is having can be seen in the following statistics. An estimated 6.6 million people in low and middle income countries were receiving HIV treatment at the end of 2010 (an increase of more than 1.35 million over 2009). As a consequence of expanded treatment, AIDS-related deaths are decreasing, and growing numbers of people with HIV are living longer and productive lives. The number of people dying from AIDS-related causes fell to 1.8 million in 2010 (down from a peak of 2.2 million in the mid-2000s). A total of 2.5 million AIDS-related deaths have been averted since 1995 due to Antiretroviral therapy (ART) being introduced, according to new calculations by UNAIDS [UNAIDS 2011].

1.2 Review of hospital data by an HIV DMP

HIV disease management (DMP) is one of the strategies used in the fight against the HIV epidemic. The core function of an HIV DMP is to increase the access to ART and do the clinical intervention when this is required. From the statistics above, it can be seen that access to treatment does make a significant impact in the fight against HIV. In South Africa, we have huge challenges in this fight and need to use our programs effectively. The HIV DMP needs an integrated approach to HIV disease management and for this reason reviewing hospital data needs to be included in the approach. The use of hospital data in an HIV DMP is standard for reporting and risk analysis purposes. The most important hospital diagnoses that would impact on the patients HIV disease management are outlined below.

1.2.1 Opportunistic infections

Opportunistic infections (OI's) are the specific range of conditions which are the main causes of mortality in HIV infected patients. They are listed as Clinical Stage IV (AIDS) by the WHO [World Health Organization 2007] and as one of the criteria to start Antiretroviral treatment (ART) [SA HIV Clinicians Society 2008]. Numerous studies have shown that the main reasons for hospital admission for HIV patients are various OI's. The intervention by the HIV DMP would be to ensure that the patient is on preventive treatment for the condition, as well as on ART. The intervention also includes checking if the patient requires treatment for the OI condition or the treatment for the OI has a drug interaction with a patient's ART regimen. Should there be a drug interaction, a review needs to be done to check if an amendment in ART dose or regimen is needed. This is to ensure that the ART remains effective. Sub-therapeutic doses will result in drug resistance and ineffective response to therapy

1.2.2 Recurrent hospital admissions for specific conditions

When a patient has recurrent hospital admissions for specific conditions, the intervention by the HIV DMP would be to check if the patient is not adhering to ART. The reason for this is that it has been shown that 80% adherence or greater is associated with lowest death rates [Regensberg et al. 2010].

1.2.3 An adverse drug reaction or toxicity related to ART

When the patient is hospitalized for an adverse drug reaction or toxicity related to ART, the intervention by the HIV DMP would be to review and if necessary change the ART medication.

For these reasons (1.2.1, 1.2.2 and 1.2.3 above), it is clear that adjudication of hospital data is critically important in HIV disease management.

1.3 Objective

In the HIV DMP that was reviewed, hospital data is currently being adjudicated manually by clinical experts. This is both time consuming and is not the most efficient use of highly-skilled resources. This thesis evaluates the electronic adjudication of hospital data for HIV patients using an expert system. An existing adjudication process is analysed, and a rule-based system to automate the process is described along with its evaluation by experts and end-users. Adjudication of hospital data using an expert system is expected to improve clinical interventions. The expert system is named *HIV-expert system*. The focus is on Opportunistic Infection (1.2.1 above) as the hospital diagnosis. The reason for this is the OI's are the leading cause of mortality in HIV patients [Groenewald et al. 2005a] and [Groenewald et al. 2005b; Giarratano et al. 2000; Zwi et al. 2000]

The process followed is that hospital diagnosis and medication information is received and this information is then processed by the *HIV-expert system*. The *HIV-expert system* applies rules which are made up of rule criteria and rule actions. The rule properties are yes/no statements and non-overlapping. The rule criteria are checked and once these are matched for rule, the rule actions are applied to the patient record. This is explained in more detail in future chapters. For a user to review the outcome of rule processing by the *HIV-expert system*, an interface in the form a dynamic website has been developed.

How the process works:

- Hospital diagnosis and medication information received
- Rule processing by the *HIV-expert system*
- Interventionist (pharmacist) checks results of rule processing on the website
- Interventionist contacts the doctor to discuss clinical intervention and agree on an action plan

The *HIV-expert system* will run in the HIV DMP and the users will be interventionist pharmacists.

It is the objective of any thesis to make a contribution to the field which is studied. The main contribution of this thesis is that it is because of the success of the electronic adjudication of pathology results by an HIV DMP, the method was extended to include hospitalization. The electronic adjudication of pathology results by an HIV DMP was presented at 2 separate conferences [van Huyssteen et al. 2003, van Huyssteen et al. 2004]. The first presentation addressed "the need to optimize resource utilisation". The findings was very positive in that "82% required no human action and could be fully processed by the system, and of these, all files monitoring adverse effects were fully automated, 18% of update files required manual intervention and turnaround time reduced from 7-14 days to 1 day" [van Huyssteen et al. 2003]. The second presentation addressed the more complex adjudication of Cd4 & viral load results. The study concluded that electronic adjudication of disease progression and effectiveness of therapy improves operational efficiencies and can be done accurately [van Huyssteen et al. 2004]. The author was fortunate to work on these 2 presentations. The success of that electronic adjudication has to a large extent provided the impetus to extend it to include hospitalization data.

One of the detailed contributions of this thesis is that a prototype website was created in the context of this thesis to evaluate accuracy and performance when the *HIV-expert system* is used by end users. While the user interface is beyond the scope of this thesis, a prototype website was built for users to get an idea of how the details in the *HIV-expert system* can be viewed.

The research methodology used for testing the HIV-expert system was a reiteration of evaluation goals. The evaluation goals are: (1) evaluation of rules, (2) evaluation of rule adjudication for sample of patients and (3) evaluation of rule creation. Qualitative evaluation was done the evaluation of rules and the evaluation of rule creation. Quantitative evaluation was done on the evaluation of rule adjudication for sample of patients.

1.4 Scope

A system to adjudicate hospital data will require a component to import hospital data. The mechanism chosen for data import, and user interface design of an adjudication system, are beyond the scope of this thesis. These aspects are incorporated in the prototype for the sake of completeness only.

1.5 Organization of this thesis

Chapter 2 is a literature review and chapter 3 is an overview of the current adjudication system (EPS). Chapter 4 starts with a high level overview of the HIV-expert system and then continues with details of the system design. Chapter 5 describes implementation of the *HIV-expert system* in SQL and chapter 7 explains how the *HIV-expert system* was evaluated. Chapters 8 and 9 cover the main results and conclusions.

Chapter 2: Background

2.1 Introduction

Automated adjudications of hospital data in an HIV DMP requires an intelligent system and hence falls into the area of Artificial Intelligence (AI). Automated adjudication is only possible because of expertise in this field and because the knowledge can be easily expressed in natural language, hence an expert system is required. Soni notes that: "Assisting physicians in making diagnosis and treatment recommendations is the most commonly found application of expert systems in medical science" [Soni et al. 2011].

This chapter starts with a definition of artificial intelligence and this is followed by the definition of an expert system, its basic structure, the development life cycle, its characteristics which include the advantages and disadvantages of using expert systems. A specific type of expert system is then discussed which is the production rule-based expert system. The last section discusses 2 areas of HIV medicine (clinical conditions in HIV and preventive therapy with co-trimoxazole) which is important for this project.

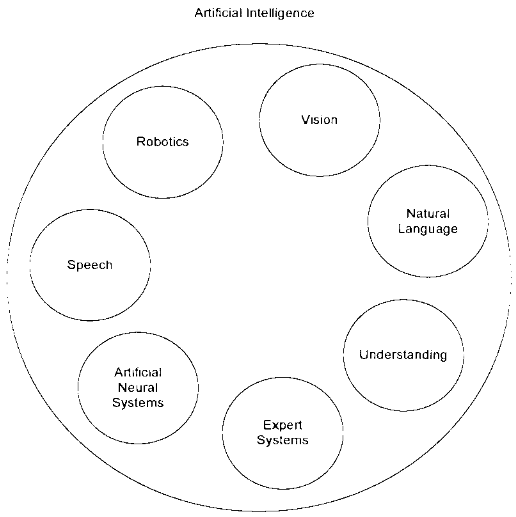
2.2 Artificial intelligence

In the Oxford Dictionary of Computing, Artificial intelligence (AI) is defined as: "a discipline concerned with the building of computer programs that performs tasks requiring intelligence when done by humans" [Illingworth 1996]. Within AI, examples of the kinds of tasks which use this are: game playing, automated reasoning, machine learning, natural language understanding, planning, speech understanding and theorem proving. Tasks can be divided into 2 groups: *intellectual* (e.g. game playing and theorem proving) and *perceptual* (e.g. hearing or seeing) [Illingworth 1996]. Expert systems and robotics are areas of AI where computer programs are built to solve problems for technological applications.

Giarratano and Riley [1989] also illustrates that AI has many areas of interest. Figure 1 is from this text. The area of medicine specifically has been significantly impacted by AI applications which have led to the construction of intelligent machines. Expert systems in medicine and the advantages and disadvantages are discussed in more detail in the latter part of this chapter.

The list of primary (and overlapping) areas into which present data artificial intelligence research can be divided broadly was originally listed by Weld [1995] and a summary of each area is provided by Doyle and Dean [1996]. They define the areas as: "(1) knowledge representation and articulation, (2) learning and adaptation, (3) deliberation, planning, and acting, (4) speech and language processing, (5) image understanding and synthesis, (6) manipulation and locomotion, (7) autonomous agents and robots, (8) multiagent systems, (9) cognitive modelling and (10) mathematical foundations. In the author's view, expert systems fall into the area of "deliberation, planning, and acting". This is because according to Doyle and Dean [1996], this area concerns methods for making decisions, constructing plans or designs to achieve specified goals, and monitoring, interpreting, diagnosing, and correcting the performance of the plans and implementations of the designs.

SOME AREAS OF ARTIFICIAL INTELLIGENCE



Extract from page 2 in GIARRATANO, JC. AND RILEY, G. 1989. Expert systems: principles and programming. PWS-Kent, Boston.

Figure 1: Some areas of artificial intelligence

2.3 Expert systems

"A good definition of an expert system is that it is a group of computer programs, along with knowledge, information and databases which act together to simulate the problem-solving and decision-making processes of a human expert within a relatively narrow domain" [Alberico and Micco 1990] . Giarratano and Riley [1989] say "An expert is a person who has expertise in a certain area. That is, the expert has knowledge or special skills that are not known or available to most people." A more technical definition from Hunt [1986] is that an expert system is "computer program that contains both declarative knowledge (facts about objects, events and situations) and procedural knowledge (information about courses of action). Its purpose is to emulate the reasoning process of human experts in a particular domain".

2.4 Basic structure of an expert system

The basic structure of an expert system includes a knowledge base, an inference engine and a user interface [Hunt 1986, Grosan and Abrahams 2011]. *The knowledge database* contains facts and heuristic knowledge. *The inference engine* contains problem solving logic. The *user interface* should be as "natural" as possible by: 1) employing language as close as possible to ordinary language and 2) understanding and displaying images. This is all at speeds that are comfortable and natural to the user [Hunt 1986]. In Figure 2, the basic steps are defined as follows: expert knowledge is used to create the knowledge database, knowledge database passes "expert knowledge" to the inference engine, the user interface passes "facts and data" to the inference engine and the inference engine passes "expertise" back to the user interface.

Hunt explains that facts and heuristic knowledge are collected in a database consisting of information about objects and events on which the knowledge base will work to achieve the desired advice. Some systems use a relational database in which the relationships between objects and events are stored explicitly for flexibility of storage and retrieval [Hunt 1986]. The knowledge base: 1) is used for organizing, controlling, propagating and updating stored knowledge and 2) it initiates searches for knowledge relevant to the line of reasoning upon which the inference system is working. Knowledge is the major factor in the performance of an expert system. This knowledge is in 2 forms: common facts and heuristics. Common facts are knowledge that is widely shared and accepted by professionals in the field. Heuristics are knowledge of good judgement and common practice or "rules of thumb" [Hunt 1986] . A traditional database is "just a database" whereas a knowledge base includes facts, assumptions, beliefs, heuristics, "expertise" and methods of dealing with the database to achieve desired results [Hunt 1986]. Reasoning techniques which manipulate the knowledge are relatively simple. It is the sophistication of the knowledge itself that is important [Ford 1991].

The inference engine: provides a process by which the lines of reasoning are formed; for example, syllogisms and other common ways of reasoning step by step from premises. In the real world knowledge and data are often inexact. Therefore some problem solving inference procedures can use degrees of uncertainty in their inference making.

The knowledge database in an expert system is developed through careful analysis of the knowledge from the "experts" in a field. The most difficult types of knowledge to obtain are: experimental, judgmental knowledge, the knowledge underlying expertise and rules of thumb and heuristic knowledge. Knowledge engineers who study Artificial Intelligence and know how to present knowledge in a computer are needed to: develop the knowledge acquisition process, to create reasoning programs to utilize the knowledge and to assure logical collection of "expert knowledge" for the creation of an effective knowledge data base. That is called "knowledge engineering" defined as the art of designing and building expert systems and knowledge based programs [Hunt 1986].

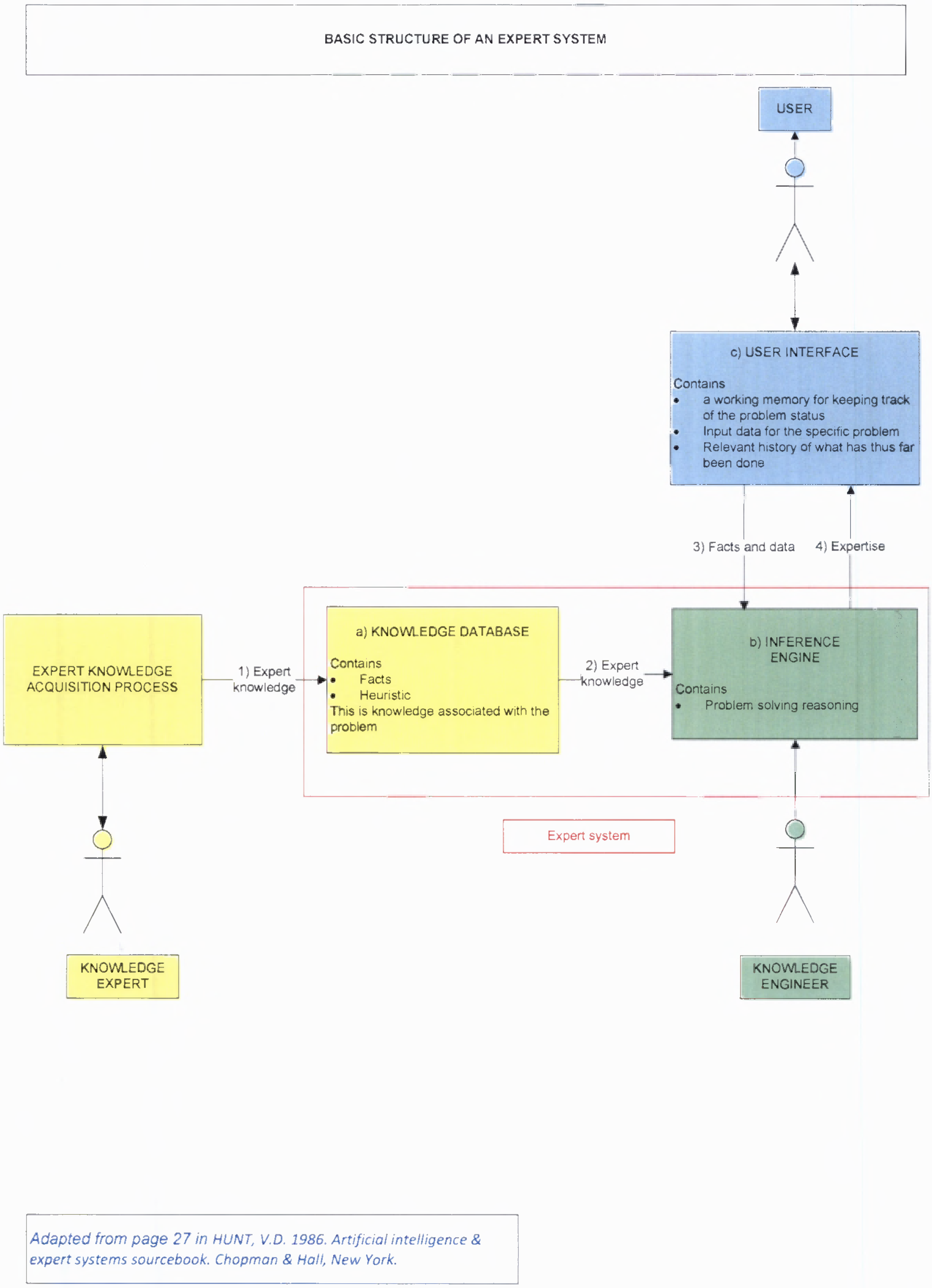


Figure 2: Basic structure of an expert system

According to Black [1986] the traditional life cycle model with the emphasis on the pre-implementations stages of "investigation, analysis, specification and design" describes satisfactorily what needs to be in place in an expert system. The process of building an expert system is called *knowledge engineering* and is done by a *knowledge engineer*. It is interesting to refer to Feigenbaum's [1982] definition: "The knowledge engineer practices the art of bringing the principles and tools of AI research to bear on difficult application problems requiring expert knowledge for their solution. The technical issues of acquiring this knowledge, representing it and using it appropriately to construct and explain lines of reasoning are important problems in the design of knowledge-based system"

Stages in the development of an expert system appear in Giarratano and Riley [1989] as: "(1) the knowledge engineer first establishes a dialog with the human expert in order to elicit the expert's knowledge. This stage is analogous to a system designer in conventional programming discussing the system requirements with a client for whom the program will be constructed. (2) The knowledge engineer then codes the knowledge explicitly in the knowledge base. (3) The expert then evaluates the expert system and gives a critique to the knowledge engineer." This process iterates until the system performance is judged to be satisfactory by the expert.

2.5 Factors affecting the performance of an expert system

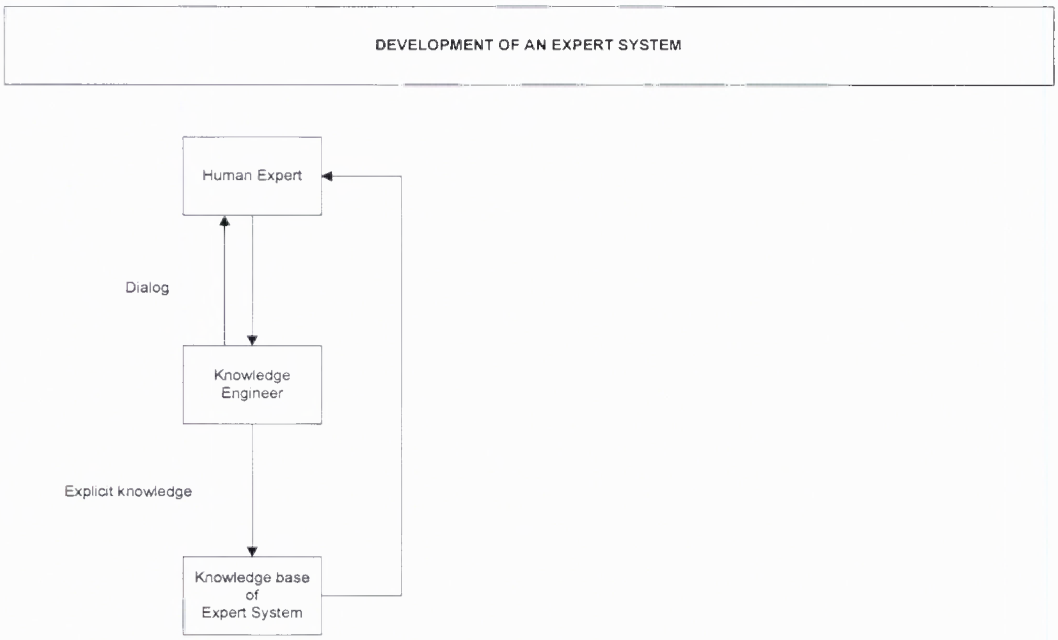
The *pre-requisites* of an expert system are listed by Hunt [1986] as: "there must be at least 1 human expert acknowledged to perform the task well, the primary source of the expert's exceptional performance must be special, the expert must be able to explain the special knowledge and experience and the methods used to apply them to a particular problem, the task must have a well-bounded domain of application". The author is in a fortunate position to be a knowledge expert and the knowledge engineer in the *HIV-expert system*. In this project, the reliability on an additional human expert (first pre-requisite 1 above) is slightly decreased.

Knowledge engineers believe that a good expert system application has these characteristics: it does not require common sense to solve, an expert will need a few minutes to a few hours to solve it and it has an expert committed to system support. Excellent quality advice must be given in a time similar to that required by an expert, or sooner. The system must be reliable, robust and permit adding, changing and removing knowledge [Giarratano and Riley 1989]. Features characterizing an expert system include the abilities to: cope with data which maybe uncertain or partial come to uncertain and where necessary multiple conclusions and explain why it is asking a particular question and how it reached a particular conclusion [Ford 1991].

2.6 Expert system *versus* conventional programming

One important aspect is that the complexity is reduced by focussing the search using rule-based heuristics. The rule-based system is able to reason about its own search effort and reasoning about the problem domain [Hunt 1986].

An expert system differs from conventional programming in the follow respects: an algorithm is used in conventional computing *versus* heuristics and inference used in expert systems; and the separation of knowledge that exists in expert systems. Expert system generally address problems having no algorithmic solution When an algorithm is contrasted with heuristics, an algorithm is defined as a well-understood procedure that is guaranteed to find a solution if it exists or to determine that no solution exists [Hunt 1986].



Extract from page 7 in GIARRATANO, JC. AND RILEY, G. 1989. Expert systems: principles and programming. PWS-Kent, Boston.

Figure 3: Development of an expert system

Heuristics are rules of thumb used by human experts. Many intelligent systems can cope with complex decision making which so far have defied algorithmic systems. These systems often entail uncertain and "fuzzy" data and highly complex chains of reasoning. In other cases, it may be relatively easy to devise an algorithm, but applying the algorithm would take too much time [Ford 1991]. Heuristics (intelligent short cuts) find solution paths without exhaustively trying each possible one. Heuristics are associated with knowledge that is informally defined and empirical, and handles