THE CITY UNIVERSITY

LONDON EC1V OHB

DEPARTMENT OF SYSTEMS SCIENCE

THESIS SUBMITTED FOR

THE AWARD OF THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN SYSTEMS SCIENCE WITH SPECIALISATION IN:

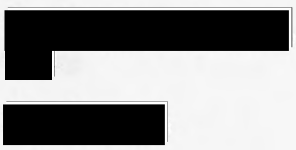
MEASUREMENT AND INFORMATION IN MEDICINE

ARTIFICIAL INTELLIGENCE AND PHYSIOLOGICAL MODELS IN MEDICINE: A PROTOTYPICAL APPROACH.

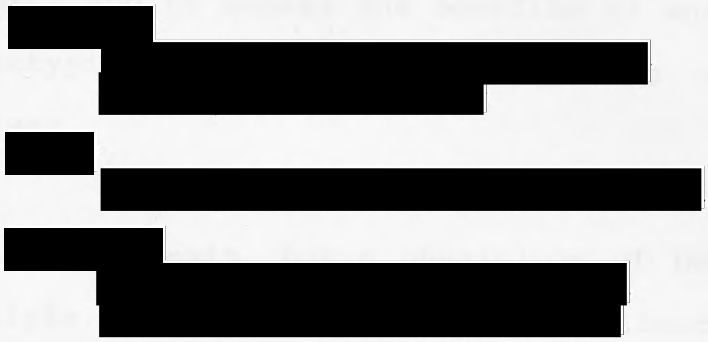
BY

AFSHIN SHAMSOLMAALI

MAY 1988



Appointments



(Warren McCulloch, 1919)

ABSTRACT

The objectives of this project were to develop a knowledge-based system coupled to a mathematical model for interpretation of laboratory data in an Intensive Care Unit; and to assess the benefits of such a coupling. A prototypical approach was taken to achieve these objectives.

In this thesis, basic physiology of body fluids and electrolyte is presented and several clinically oriented knowledge-based systems are reviewed. The two versions of the developed prototype are described, the problems associated with coupling symbolic programs to numerical programs are emphasised and a description of the methodology adopted is presented, as are simulation results of clinical cases. A critical discussion of the knowledge representation and reasoning methodology is given, and contributions to systems science and medical informatics are highlighted.

ACKNOWLEDGEMENTS

I would like to express my gratitude to those who made this work possible and who created a stimulating environment to work in.

Professor Ewart R. Carson for his supervision and excellent handling of administrative difficulties.

Dr. Derek G. Cramp for his supervision and his unprecedented support and friendship during the past three years.

[redacted] for his learned input of clinical expertise.

[redacted] for supplying Macpee and for his support during the course of development of the interface.

[redacted] whose contribution to this work and to my progress was invaluable.

[redacted] for supplying ideas and suggestions when they were most needed.

[redacted] for his invaluable help with clinical cases.

I would also like to thank colleagues at the departments of Chemical Pathology and Medical Informatics, Royal Free Hospital, for their support and friendship during the past two years. These were: [redacted]

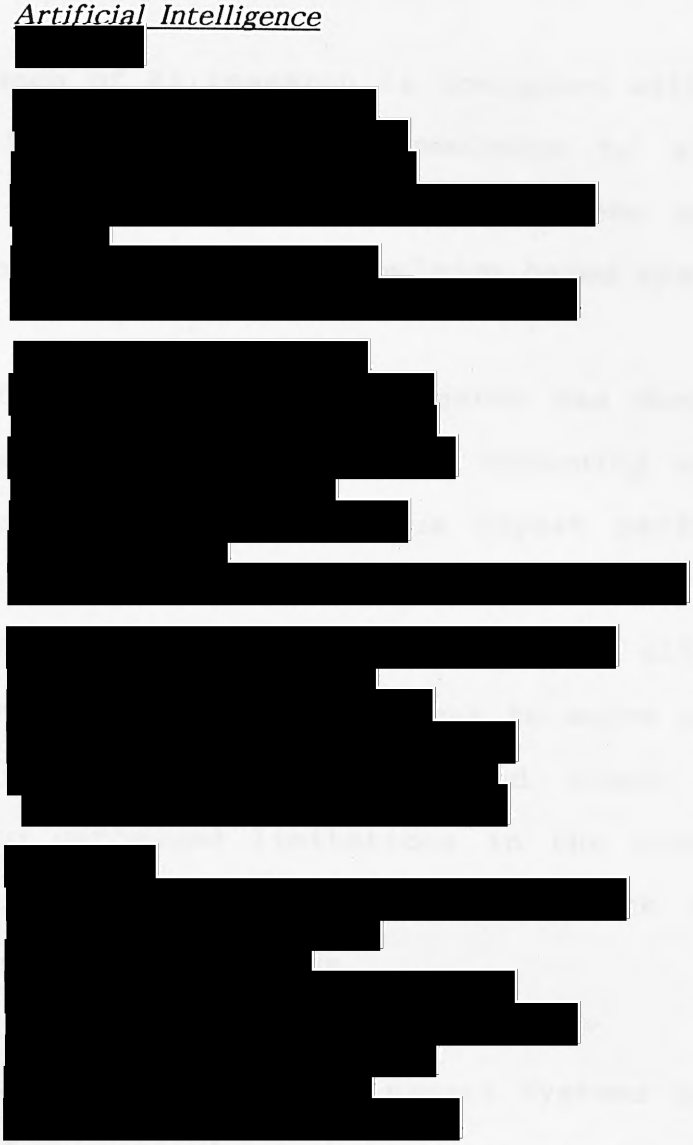
Finally, I would like to thank [redacted] for her encouragement and understanding during my postgraduate studies.

This work was supported by a grant from the Department of Health and Social Security (DHSS).

TABLE OF CONTENTS

CHAPTER 1

Artificial Intelligence



(Adrienne Rich, 1961)

1.0 INTRODUCTION

Since World War II, computer scientists have tried to develop techniques that would allow computers to act more like humans, that is intelligently. The entire research effort, including decision making systems, robotic

devices, and various approaches to computer speech and vision, is usually called Artificial Intelligence (AI).

A branch of AI research is concerned with developing programs that use symbolic knowledge to simulate the behaviour of human experts. These systems are commonly known as ***expert systems*** or ***knowledge-based systems***.

The first period of AI research was dominated by a naive belief that a few laws of reasoning coupled with powerful computers would produce expert performance. As experience accrued, the severely limited power of general-purpose problem-solving strategies ultimately led to the view that they were too weak to solve most complex problems (Newell, 1969; Newell and Simon, 1963). In reaction to perceived limitations in the overly general strategies, many researchers began to work on narrowly defined application problems.

By the mid-1970s several expert systems had begun to emerge. A few researchers who recognised the central role of knowledge in these systems then initiated efforts to develop comprehensive knowledge representation theories and associated general-purpose systems (Bobrow and Winograd, 1977; Minsky, 1975; Szolovits, Hawkinson and Martin, 1977). Within a few years it became apparent that these efforts had limited success for reasons similar to those that doomed the first general-purpose problem-solvers. "Knowledge" as a target of study is too broad

and diverse; efforts to solve knowledge-base problems in general were premature. On the other hand, several different approaches to knowledge representation proved sufficient for the expert systems that employed them. The lesson learned from these experiences was referred to in what Professor Edward Feigenbaum of Stanford University, in an invited paper, stated:

> *"The performance level of an expert system is primarily a function of the size and the quality of a knowledge base it possesses."*
>
> *Feigenbaum, 1977*

In short, an expert's knowledge provides the key to expert performance, while knowledge representation and inference schemes provide the mechanism for its use.

Experts perform well because they have a large amount of compiled, domain-specific knowledge stored in long-term memory. Compiled knowledge takes two forms: first principles and general theories on one hand, and heuristics and domain theories on the other. The amount of knowledge an expert requires is such that it is nearly impossible to gain it all from experience.

Experts acquire knowledge of the first principles and general theories that are regarded as basic to their profession. Then they begin to practice their profession. In the process, experience is gained and knowledge is

recompiled. Experts move from a descriptive view of their profession to a procedural view. Practising experts hardly ever explain their recommendations in terms of first principles or general theories. If they encounter unusual or complex problems, however, they will return to first principles to develop an appropriate strategy.

A knowledge-based system should therefore, possess similar knowledge of heuristics and first principles in order to simulate the decision making process of an expert.

1.1 PHYSIOLOGICAL MODELS

Models are always used, consciously or not, by experts. Clinicians have a conceptual model of their domain of expertise which they use for decision-making, prediction and explanation. This model is formed gradually in the process of knowledge acquisition and accumulating experience. Quite often, such models are incomplete due to lack of full understanding of the mechanisms of action of the domain.

Various attempts have been made to code this kind of conceptual model by using techniques like decision trees, flow charts, mathematical simulations and more recently causal networks. All these techniques have strengths and weaknesses.

Decision trees simulate the decision-making process by creating paths that correspond to the expected observations before a decision (diagnosis) is reached. Alternative paths are tried to establish a decision. The decision-making process is pattern matching. This technique is very efficient if the represented domain is well-defined and the level of uncertainty is low.

Mathematical simulation has been in use for a number of years to develop models to study various aspects of represented domain. Mathematical modelling of physiological systems has received substantial interest in the recent years (Guyton, Coleman, and Granger, 1972; Carson, Cobelli and Finkelstein, 1983; Finkelstein and Carson, 1985; Cramp, 1975).

Knowledge is represented and encapsulated by means of mathematical relations. Both first principles and emprical knowledge are represented. Parameters of such models are adjusted (fitted to data) to achieve realistic simulations. Such models are a rich source of knowledge; the represented knowledge, however, is implicit by nature.

With the introduction of AI techniques, it became possible to represent knowledge symbolically. In the early years only associational knowledge was represented symbolically (Shortliffe, 1976; Pople, 1977; Weiss, Kulikowski, Amarel and Safir, 1978). There was then a

move towards the representation of causal reasoning by using causal networks and qualitative simulation (Patil, 1981; Kuipers, 1984). Knowledge represented in this way is explicit and therefore readily accessible for explanation, prediction and so forth. Physiological systems, however, cannot be completely represented in this way because of the uncertainties involved in understanding the mechanics.

Generally speaking, there is no perfect method for representing physiological systems. The methods described have their strengths and weaknesses. Natural selection dictates that the systems combining different techniques of knowledge representation and utilising their strengths will survive and evolve.

1.2 SCOPE OF THE PROJECT

The objective of the present research programme was to develop a prototype knowledge-based system coupled to a mathematical simulation to be used in a clinical setting. For the research group involved, it was an experiment to assess the problems (both technical and conceptual) of such coupling.

Coupling essentially numerical algorithms with components of symbolic computing has attracted a certain amount of attention (Kowalik, 1986; Kowalik, Chalfan, Marcus and Skillman, 1986). This attention, however, is not from a uniform point of view. The "conventional"

programmer, on one hand considers symbolic computing as a vehicle for building "intelligent" front-ends to their well developed algorithms, and hence make them more usable and accessible. The "symbolic" programmer on the other hand, wish to use well-developed algorithms and simulations to solve aspects of the problem at hand. In the latter case, the level of coupling varies according to the requirements and objectives. One could have a "shallow" coupling where the simulation and the knowledge base view each other as "black boxes" with parameters being passed from one to the other. A "deep" coupling on the other hand, requires more interaction between the two sources. Symbolic and numeric representations are based on a similar conceptual understanding of the domain; the challenge is how to combine the two in such a way so as to reduce the size of the knowledge base and at the same time enhance the quality of the knowledge represented. This implies that the two knowledge sources must be complementary to each other at every level, ie. representation and computation. This is by no means a trivial problem and is still the subject of much research.

A "deep" coupling may not always be desirable as the implementation of these systems is still some way away. It must be pointed out, that this is not a "black and white" situation, where a system is either "deeply" or "shallowly" coupled to a numerical simulation or algorithm; there are all shades in between.

The system described here is essentially coupled in a "shallow" manner, for reasons which will be described in chapters 4 and 6.

The domain of interest is fluid and electrolyte disorders in the Intensive Care Unit environment. This is an area where laboratory and clinical (bedside) data are used extensively to assess the patient's condition and subsequent fluid therapy. Application of computer aided decision making in this environment is highly desirable as a vast number of variables are measured routinely several times a day. This complicates the task of decision making for clinicians.

On the surface this seemed to be an ideal environment, both because of complexity of decision making and because of the fact that several mathematical models were available in this domain (Dickinson, Ingram and Ahmed, 1987; Ikeda, Marumo, Shirtataka and Sato, 1979; Flood,Carson and Cramp, 1986).

The model used was Macpee (Dickinson, Ingram and Ahmed, 1987) which is a model of circulatory and fluid electrolyte metabolism. Macpee is able to simulate physiological and pathophysiological disturbances over time. Hence time can be represented in an abstract manner, in the knowledge based system. Because of the constraints which will be described, time is only represented at the therapeutic level.

The prototype system provides a diagnosis and suggests treatment regimens for patients with fluid and electrolyte disorders using the data input to the system. Subsequent justification of the therapeutic measures are carried out by passing the instructions to MACPEE and simulating for an appropriate period of time.

The model therefore, does not play any role in diagnosis or treatment generation; it is used as a tool to justify the suggestions of the treatment module. Facilities are provided to enable the user to simulate alternative therapeutic measures.

1.3 SCOPE OF THE THESIS

This thesis is centred upon the developed prototype for interpretation of the laboratory data. The development phases and the coupling technique are described.

Chapter 2 describes the physiology and the pathophysiology behind the represented knowledge-base.

Chapter 3 reviews a sellection of the clinically-oriented knowledge-based systems. The systems reviewed incorporate within them some kind of physiological model of the underlying mechanism. Problems associated with each system are discussed and their strength and novelty highlighted. In addition, the work of colleagues at the Royal Free Hospital, on the development of MODEL, a tool

for rapid development of physiological models from conceptual description, is reviewed.

Chapter 4 describes the structure and the development phases of two versions of the prototype in detail.

In *Chapter 5* the reasoning mechanisms of the two versions are illustrated by means of four examples and the responses are compared and contrasted.

In *Chapter 6* a critical discussion of the two versions developed is given, together with suggestions for improvements and suggestion for future work that would allow complete integration of numerical simulation techniques with a knowledge-based system.

Chapter 7, highlights the contributions made both to systems science and medical informatics.

Appendix 1 contains some examples of the output windows.

Appendix 2 gives a listing of the developed program in LPA-Prolog.

CHAPTER 2

2.0 BODY FLUID AND RELATED DISORDERS

The constancy of the fluid surrounding the body cells is an essential requirement for their proper function. Three major variables must be accurately controlled: volume, concentration (total osmolality), and pH (hydrogen ion concentration).

The majority of patients admitted to an intensive care unit suffer from abnormalities of salt and water balance and circulation and since the control of all three is closely interrelated such disturbances may constitute a threat to life and therefore need urgent correction. In this chapter the pathophysiological mechanisms of some of these disorders are described.

2.1 BODY WATER

2.1.1 Volume

The distribution of body water is best related to the "lean body mass" (LBM). This is the total body mass exclusive of fat, since fat takes no immediate part in exchange of water. The contribution of fat to the body weight varies widely between different individuals, while the LBM is quite closely and constantly related to height. Fat can of course release much water when metabolized, as in the familiar example of the camel's hump, but this is a slow process and therefore not

relevant to acute clinical situations. Fat usually comprises some 10-20 per cent of the total body weight.

The lean body mass is composed of about 30% solids and 70% water. The water is in two major compartments, the intracellular fluid (50% LBM) and the extracellular fluid (20% LBM). The extracellular fluid (ECF) is again divided into two compartments, the plasma (5% LBM) and the intercellular fluid (15% LBM), (see fig.2.1). The intercellular fluid bathes the cells and supports them, and can also further be considered as comprising several compartments. These are considered in more detail in the section on the extracellular fluid.

The plasma compartment is the only one in direct communication with the environment, through the gut, lungs, kidney etc., and most of the reflex mechanisms controlling homeostasis operate through it. It is also the only one which can be readily and repeatedly sampled, and it is therefore on the basis of changes in the composition of the blood plasma that deductions are made as to the state of the body fluids. While such deductions are usually valid, and of the greatest value in patient care, there are occasions when they seem to mislead, as when a high serum potassium is associated, apparently anomalously, with a low intracellular concentration of the ion. A knowledge of such apparent anomalies is essential in the handling of individual patients.

Plasma 5% (3 kg)	Intercellular 15% (9 kg)	Intracellular 50% (30kg)

Figure 2.1- Composition of body fluid (of 60 kg fat-free weight).

2.1.2 Concentration

The movement of water across the plasma membrane of a cell is governed by the relative osmotic pressure of the intracellular and extracellular fluids. Cells are hyperosmolar relative to the plasma and therefore they will tend continually to take up water which they must then actively excrete. Osmolality is therefore the appropriate unit of measurement of concentration in relation to water equilibrium. In contrast, in the control of pH it is the equilibrium of the ionized constituents of the fluids that must be maintained, and therefore at neutral pH the sum of the anionic molecules must balance that of the cationic molecules. The electrolyte concentrations of the compartments differ very markedly (table 2.1) and these differences are actively maintained in a state of dynamic equilibrium. While each ionized particle contributes to the osmotic pressure, unionized substances such as glucose and urea may also contribute substantially to the total osmolality of the body fluids. Such substances can be of the greatest importance in relation to the bulk flow of the fluid into and out of the cells; they also contribute to the osmolality of glomerular filtrate and so affect the capacity of the kidney to produce a concentrated urine.

2.1.3 Water Balance

The normal (70kg) adult (60kg lean body mass) drinks about 2000ml of water daily and obtains a further 500ml from the oxidation of the food. Total body water is approximately 42000ml. Losses from body are due to

Major Constituents of Blood Plasma

	mmol/litre
Cations	
Sodium	135-142
Potassium	3.5-4.5
Calcium	2.4-2.6
Magnesium	0.7-0.9
Anions	
Bicarbonate	23-30
Chloride	98-107
Phosphate	0.75-1.35
Protein	depends on their isoelectric points

Major Constituents of Intracellular Fluid

	mmol/litre
Cations	
Sodium	10
Potassium	160
Magnesium	12
Anions	
Bicarbonate	8
Chloride	2
Phosphate(organic)	45
Protein	Partially ionised and act as buffer anions

Table 2.1- Table of major constituents of blood plasma and intercellular fluid.

evaporation from the skin, lungs, sweat etc., faecal losses and urinary excretion. The minimal urine volume required to excrete the normal waste products is 500ml, and at this volume the urine will be at the maximum osmolality of which the kidney is capable of sustaining. The osmolar concentration of the urine excreted is controlled by the antidiuretic hormone (ADH) which is secreted by the posterior pituitary. Increasing the concentration of sodium ion (Na^+) in the blood perfusing the internal carotid artery causes more hormone to be secreted and consequently more water is reabsorbed by the distal tubules. The result is a concentrated urine and dilution of the plasma.

2.1.4 Syndromes in Disease

a) Dehydration

Loss of water in excess of intake results in dehydration which is manifested as a generalized diminution of the volume of all the body compartments in proportion to their relative volumes. Infants have a much smaller total volume of body water relative to intake and output, and a higher proportion in the ECF. They are therefore especially susceptible to rapid dehydration. Dehydration can follow deficient intake or loss of fluid by any one of a variety of routes. Intake insufficient to replace inevitable loss is the most obvious cause of dehydration and this can be aggravated by concomitant insensible losses of water and salt in hot climates. While losses of water relatively free of

electrolytes may occur, more usually there are accompanying electrolyte losses which depend on the origin of the fluid lost. Damage to the skin results in large losses of plasma-like fluid, rich in protein and extracellular electrolytes. Large volumes of fluid are secreted daily into the gastrointestinal tract and if these are not reabsorbed rapid dehydration is inevitable. Renal insufficiency can lead to excessive losses of water. As renal failure advances, the capacity of the kidney to concentrate diminishes, and the loss of a urine of low osmolality may lead to dehydration. Defective ADH production leads to the most profound diabetes insipidus, with uncontrollable thirst and rapid dehydration.

b) Overhydration

This can result from inappropriate secretion of ADH or ADH-like peptides. When an excess of water is retained it is distributed throughout the body compartments. Serum sodium may fall to very low levels (less than the 120mmol/l), but it is the accompanying cellular overhydration that is damaging, leading to coma, convulsions and death. The capacity of the normal kidney for excreting excess water is, however, so great that overhydration from excessive drinking is almost impossible, but if the kidney function is severely damaged, overhydration is a real possibility due to failure of excretion. In the special instance of the first 24-48 hours following surgery there is excess of ADH production, with a fall in serum sodium. Misguided attempts to correct this with saline solutions (often

hypotonic to blood plasma) can lead to rapid overhydration and pulmonary oedema.

2.2 INTRACELLULAR FLUID (ICF)

2.2.1 Control of ICF

The ICF is separated from the ECF by the cell's plasma membrane. This membrane contains phospholipids and is therefore readily permeable to lipid-soluble ions and molecules, such as Na^+, K^+ and urea. In addition, active absorption of non-lipid, water-soluble substances can occur by the process of pinocytosis or by absorption of a lipid-soluble reaction product following modification of the substance at the cell surface by enzymes. Hyperosmolarity of the ICF relative to ECF results in the continual uptake of fluid and ions by the cells, some of which are then actively excreted; e.g. selective excretion of Na^+ maintains the high intracellular concentration of K^+. Any damaging influence on the cell such as anoxaemia or changes in hydrogen ion concentration (H^+) will lead to a reduction in its capacity to excrete Na^+ and water. In the special instance of a fall in the intracellular pH there is a marked tendency to lose K^+ from the cells.

In all tissues the maintenance of the electrical potential across the cell membrane, which reflects the differential ionic concentrations on the two sides, is essential to the proper functioning of the cell. In addition to K^+, the divalent ions Ca^{++} and Mg^{++} are of

particular importance in maintaining this potential. In contractile tissues a low concentration of Ca^{++} in the ECF leads to tetany. A high level of serum K^{+} will aggravate this effect while very low serum K^{+} levels may be associated with muscular paralysis.

2.2.2 Changes in Disease

The volume, composition and pH of the ICF must alter in any condition, local or general, which damages the cells. Unfortunately we are seldom able to measure these changes directly, even when they are established and relatively constant. When rapid changes are occurring, as when an illness is developing or regressing, measurement is even more difficult and inferences have to be made from accompanying changes in plasma. In dehydration or overhydration, changes are distributed throughout the body water, but when water loss or gain is associated with sodium gain or loss, then it is mainly the ECF which alters in volume. Large changes in the ECF volume (10-12 litres or more) can be accommodated with remarkably little adverse effect if the fluid is isotonic, while an increase of only about 4% in the ICF may lead to convulsions. If the osmolality of the ICF is to remain reasonably constant then loss of water by the cells must be followed by loss of potassium which will be excreted in the urine. Similarly rehydration of the cells must be accompanied by adequate potassium replacement if a low serum potassium is to be avoided.

There is no direct hormonal feedback control of the body potassium. The balance between cellular and intercellular fluids is maintained by the cells retaining this ion and actively transferring sodium to the outside. Any change in the intracellular equilibrium is corrected by retention or loss. The uptake of potassium by the cells is markedly accelerated by the uptake of glucose under the influence of insulin. In uncontrolled diabetes mellitus intercellular (H^+) rises due to starvation, potassium is lost from the cells and excreted in the urine. When treatment with insulin and rehydration is started, the rapid uptake of potassium from the plasma can lead to a dangerous hypokalaemia.

It has recently been found that, in severe traumatic lesions, insulin is not secreted. A diabetic type of condition can thus arise with loss of sodium, potassium and water in the urine.

A proper balance of Ca^{++} and Mg^{++} is essential for the maintenance of intracellular equilibrium not only because of their immediate effect on membrane potential but also because of their effects on membrane permeability and their properties as co-enzymes.

2.3 EXTRACELLULAR FLUID (ECF)

2.3.1 Distribution

The extracellular fluid includes all fluids that are not inside cells and it is therefore a very inhomogenous

compartment. The two main sections are the plasma water (5% LBM) and the interstitial water (10-12 per cent LBM). These two are in rapid equilibrium with each other, with the cells, and with the environment. The plasma water is mainly free water but in the interstitial compartment it is bound in the dydrated gel of the connective tissue mucoprotein that surrounds and supports the cells. If it were not bound in this way the fluid would be subject to gravitational forces and would accumulate in dependent parts. Only when the gel is over-saturated does free fluid appear. This is subject to gravity and may appear as oedema.

The main difference between plasma and interstitial fluids lie in the higher protein concentration and therefore higher colloid osmotic pressure of the plasma. As a result some minor differences in the electrolyte concentrations between the two compartments are found.

2.3.2 Volume and Concentration

The volume of plasma is controlled by the dual mechanisms of ADH from the pituitary and aldosterone from the adrenal cortex. The volume of the interstitial fluid is dependent on the transcapillary forces between it and plasma. While Na^+ concentration controls the output of ADH and regulates the excretion of water by the kidney, aldosterone causes tubular reabsorption of Na^+ from the glomerular filtrate. While the ADH control is very rapid, operating within minutes of changes in plasma

concentration, the aldosterone system is much slower, operating only within hours of plasma volume changes.

2.3.3 Disorders of ECF

Changes in either the cells or the environment are usually reflected in changes in the ECF which, as mentioned earlier, can be readily sampled. Disorders of the ECF can arise in various ways: (i) disturbances of hormonal control related to ADH secretion, or aldosterone production as in Addison's disease. Such changes lead to failure of homeostatic mechanisms and disorder arises rapidly; (ii) defects of absorption or excretion mechanisms (kidney, lungs, intestine). These can be followed by limited compensation, but the capacity of the body to rectify ECF disorders from these causes is very small; (iii) excessive gains or losses from the cells (e.g. in diabetic ketoacidosis); and (iv) excessive gains or losses to the outside (e.g. following burns or diarrhoea).

2.4 SERUM SODIUM

High levels of serum sodium (hypernatraemia) are associated only with gross water dehydration and concentration of the plasma. The symptoms are those of gross dehydration. Excessive retention of sodium is more often associated with water retention, and finally the rise in the volume of ECF is clinically manifested as peripheral oedema. About 10-12 litres of excess ECF must accumulate in a 70kg man before pitting oedema can be detected. Such peripheral oedema is relatively harmless

to the patient, but if pulmonary oedema appears it can be rapidly fatal.

Low levels of serum sodium (hyponatraemia) are relatively common. Excessive sodium loss leads to a compensatory loss of ECF and to a much lesser extent of cellular water. If the plasma volume falls sufficiently the blood pressure can not be maintained, glomerular filtration rate falls and serum creatinine and urea associated with prostration and lassitude in the patient, has come to be known as the 'low sodium syndrome'. It is found not only in Addison's disease but also (along with potassium deficiency) following prolonged excessive administration of diuretic drugs.

Oversecretion of ADH causes water retention in excess of sodium and very low serum sodium levels. Deficient secretion of ADH leads to excess water loss and to hypernatraemia if thirst is not satisfied. Excess aldosterone secretion leads to sodium and water retention and potassium loss. This occurs in association with hypertension, renal failure, hypoalbuminaemia with a low plasma volume, or following trauma such as surgery or burns.

Absorption and excretion control disorders are really confined to the kidney. Absorption of sodium from the gut is seldom impaired and excessive intake induces nausea and vomitng. In renal failure the capacity of the kidney to retain or excrete sodium is often reduced, and

therefore excess intake leads to oedema whereas insufficient intake gives rise to a low sodium syndrome.

Excess loss of sodium to the outside occurs in diarrhoea, from severe burns, and during heavy sweating. The serum concentration may be raised or lowered depending on the amount of water lost relative to sodium. Excessive gains usually result from overtransfusion of saline solution, particularly if renal function is impared. Many infused solutions are hypotonic, and if given in excess are therefore usually associated with hyponatraemia. This condition is most easily detected clinically by daily weighing of the patient.

Excess loss of sodium can occur into the cells when the sodium pump is not operating; this may happen in anoxaemia, acidaemia, or intracellular glucose deficit. This, with the accompanying loss of potassium from ICF to ECF and rise in serum potassium, is sometimes called the "sick-cell syndrome".

From the considerations above, it will be clear that the level of sodium does not reflect the total body content. High levels may be associated with a gross deficit of ECF and vice versa. Estimation of both concentration and volume of distribution are necessary to assess the situation and correct it properly.

2.5 OEDEMA

Oedema means the presence of excess interstitial fluid in the tissue. Any factor that increases the interstitial fluid pressure high enough can cause excess interstitial fluid volume and therefore cause oedema.

Oedema usually is not detectable in tissues until the interstitial fluid volume has risen to about 30% above normal. In serious cases of oedematous, the interstitial fluid volume can increase to several hundred per cent above normal.

2.5.1 CAUSES OF OEDEMA

1. Abnormal Capillary Dynamics

Several different abnormalities in these dynamics can increase the tissue pressure and in turn cause extracellular fluid oedema. The different causes of extracellular fluid oedema are:

a. *Increased capillary pressure,* which causes excess filtration of fluid through the capillaries.

b. *Decreased plasma protein,* which causes reduced plasma colloid osmotic pressure and hence, failure to retain fluid in capillaries.

c. *Lymphatic obstruction,* which causes protein to accumulate in the tissue spaces and therefore causes osmosis of fluid out of the capillaries.

d. *Increased capillary permeability,* which allows leakage of excess fluid and protein into the tissue spaces.

2. Fluid Retention By Kidney

When the kidney fails to excrete adequate quantities of urine, and the person continues to drink normal amounts of water and ingest normal amounts of electrolytes, the total amount of extracellular fluid in the body increases progressively. This fluid is absorbed from the gut into the blood and elevates the capillary pressure. This in turn causes most of the fluid to pass into the interstitial fluid spaces, raising the interstitial fluid pressure. Therefore, simple retention of fluid by the kidneys can result in extensive oedema.

3. Heart Failure

Heart failure is one of the most common causes of oedema. When the heart is unable to pump blood out of the veins with ease, blood dams up in the venous system. The capillary pressure rises, and serious "cardiac oedema" occurs. In addition, the kidneys often function poorly in heart failure resulting in even more oedema.

2.6 SUMMARY

In this chapter a brief account of the physiology and pathophysiology of the body fluid and electrolyte metabolism was given. The disorders and conditions described are those represented in the developed prototype. The above discussion of disorders therefore constitutes the conceptual understanding behind the knowledge represented in the prototype which will be described in chapter 4.

CHAPTER 3

(Jose' Ortega y Gasset)

3.0 REVIEW OF EXTANT SYSTEMS

In this chapter work carried out by other researchers is reviewed. There are obviously a large number of systems that could be reviewed, however, the emphasis here is on the clinically oriented computer systems which include within them some kind of model of the underlying mechanism. Clinically oriented computer systems are reviewed because of the uncertainties involved in medicine, and therefore the method used in attempting to deal with uncertanties can be enlightining. The reviewed systems are selected because of the methodology rather than the domain they represent.

3.1 ABEL

This is the work of Ramesh Patil (1981), carried out at the Computer Science Laboratory at MIT under the supervision of Peter Szolovits and William B. Schwartz. ABEL stands for Acid Base ELectrolyte. This was pioneering work on the causal representation of physiological and pathophysiological knowledge. Because of its impact on the later work, it is reviewed in some detail, with particular emphasis on the methodology of

knowledge representation and the concept of a Patient Specific Model.

3.1.1 INTRODUCTION

The developers had three main objectives.

1. To develop a representation of causal medical knowledge.

2. To develop a case-specific "understanding" of illness. This understanding should be capable of describing subtle interactions between diseased and normal physiological mechanisms, and therapeutic interventions.

3. To develop a set of reasoning procedures to combine the aggregated phenomenological knowledge of disease associations with the detailed pathophysiological knowledge of disease processes. The phenomenological knowledge is necessary for efficient diagnostic exploration; the pathophysiological knowledge for proper understanding of a difficult case.

ABEL consists of four major components:

1. The patient specific model (PSM),
2. The Global Decision Making component,

2. The diagnostic component,

3. The therapy component.

Figure 3.1, demonstrates the relation between these components.

The PSM describes the physician's understanding of the state of the patient at any point during diagnosis and management; it is intended to be the central data structure with which other components of the system may reason.

The global decision making component is the top level program which has the responsibility of calling the other programs with specific tasks. It also modifies the patient specific model to reflect the revised state of the patient; calls upon the diagnostic and therapeutic programs. Note that at every step the global decision maker can evaluate each of the possible sets of actions and choose the most desirable one.

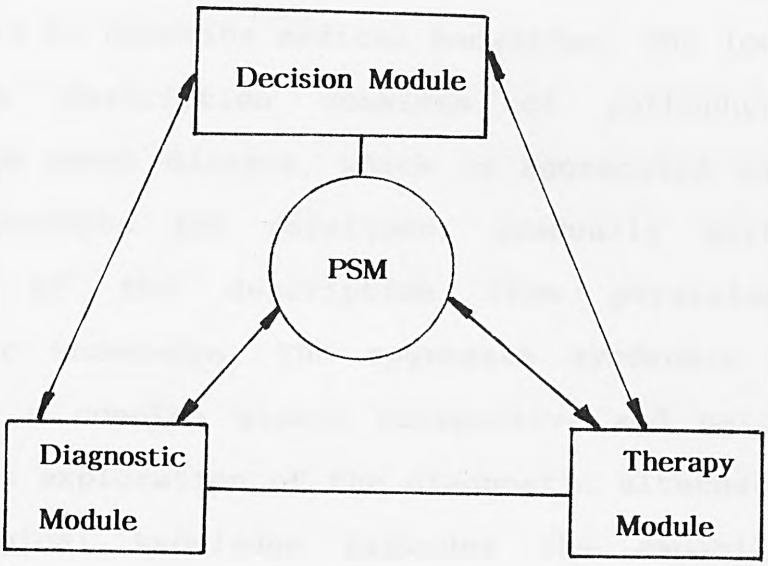


Figure 3.1- Schematic diagram of the overall structure of ABEL.

3.1.2 KNOWLEDGE REPRESENTATION

A hierarchial multi-level representation scheme is developed to describe medical knowledge. The lowest level of this description consists of pathophysiological knowledge about disease, which is aggregated into higher level concepts and relations, gradually shifting the context of the description from physiological to syndromic knowledge. The aggregate syndromic knowledge provides a concise global perspective and helps in the efficient exploration of the diagnostic alternatives. The physiological knowledge provides the capabilities of handling complex clinical situations arising in patients with multiple disturbances, evaluating the physiological validity of the diagnostic possibilities being explored, and organising a number of fragmented facts into a coherent causal description.

3.1.2.1 Anatomical Knowledge

This includes:

a) *part-of* hierarchy for organ systems;

b) *connected-to* relations that provides material flow information;

c) *contained-in* and *position* relations which provide anatomical relations between anatomical sites.

This knowledge is not currently used.

3.1.2.2 Aetiological Knowledge

Disease categories are primarily organised around the organ systems; eg. renal disease. Regardless of the cause of say renal failure, all the diseases causing renal failure share common symptoms.

In a manner similar to the anatomical categorisation, the diseases with common aetiology share symptoms common to the disease mechanism.

3.1.2.3 Physiological Knowledge

The knowledge necessary to deal with fluid, electrolyte and acid-base disorders is represented. The physiological knowledge about fluids and electrolytes in the program deals with: fluid compartments of the body and the distribution of body fluids in these compartments; the composition of fluid in each compartment; the space of distribution of solutes, exchange of fluid and electrolytes between compartments, and the homeostatic mechanisms for regulating the quantity and composition of the body fluids.

3.1.2.4 Disease Knowledge

A disease is defined in terms of its anatomical involvement, its temporal characteristics, its aetiological characterisation and its pathophysiology. As

each of the anatomical, aetiological and physiological knowledge is hierarchically organised, the locus of a disease along each of the disease definitions can then be derived from these loci. The basic medical knowledge about anatomy, aetiology etc. provide a framework for describing and organising the disease hierarchy.

3.1.2.5 Causal Link

There is a need to know how a cause relates to an effect, as well as other contextual information influencing the causal relation. To capture this information, the description of a causal link has associated with it a **multivariate relation** between attributes of the cause and the effect, the context, and the assumptions which constrain the causal relation.

3.1.2.6 Multi-Level Causal Description

The clinicians do often consider a difficult case at several levels of detail. In order to be effective, the program must be able to describe the problem briefly yet still be able to take low level details into consideration. The program's medical and case-specific knowledge is represented at five levels of detail, ranging from pathophysiological to phenomenological levels of knowledge.

Each level of description can be viewed as a semantic net describing a network of relations between diseases and findings. Each node represents a normal or

abnormal state of a physiological parameter and each link represents some relation (causal, associational, etc.) between different states.

A state is represented as a node in the causal network. Associated with each node is a set of attributes describing its temporal characteristics, severity or value etc. A node is called **primitive** if it does not contain internal structure and **composite** if it can be defined in terms of a causal network of states at the next more detailed level of description. One of the nodes at that more detailed level is designated as the **focus node** and the causal network is called the **elaboration structure** of the composite node.

Because of the fixed number of levels in the multi-level description, the program's ability to aggregate causal description is limited. To overcome this the notion of a compiled link which represents a causal pathway is introduced. The compiled link allows the selective exploration of commonly occurring causal paths more deeply than others without degrading the quality of deduction. This also provides the ability to activate nodes which are not immediate neighbours of the node under consideration.

The presence or absence of a causal relation between a pair of states can change their diagnostic and prognostic interpretations. The system has the capability

of hypothesising the presence or absence of a causal relation. This is the primary reason why links are considered objects in their own right rather than simply an ordered pair of states.

3.1.3 PATIENT SPECIFIC MODEL (PSM)

A PSM is a multi-level causal model, each level of which attempts to give an account of the program's understanding of the patient's case. Each PSM contains all the diseases and findings that have been observed or concluded in a given patient along with hypothesised diseases, findings and their interrelationships, which together form a **coherent** explanation. Within each PSM, the known and the hypothesised diseases, findings and their interrelationships are mutually complementary, while the alternate explanations which are mutually exclusive are competing to explain a patient's illness.

The PSMs are implemented using a **Patient Specific Data structure** (PSD). The PSDs are organised in a tree. The PSD in the root position of the tree contains observed findings and structure common to all the PSMs. Differing interpretations of the observed findings are described by creating inferior PSDs each containing incremental changes (additions as well as deletions) to their superior PSD. Each PSD in the tree inherits from its superior all the structure present in them except that which is explicitly deleted. The list of PSMs at any

given instant of diagnosis is called *causal hypothesis list* (CH-list).

All the new information received is always added to the root PSD, the PSD common to every PSM.

The PSMs are created and augmented using structure building operations:

Initial formulation to create the initial set of PSMs from the presenting complaints and lab results.

Aggregation to summarize the description at a given level of detail to the next more aggregate level.

Elaboration to disaggregate the description at a given level to the next more detailed level.

Projection to hypothesise associated findings and diseases suggested by states in the PSM.

Constituent summation and **decomposition** to evaluate the combined effects of multiple aetiologies and to evaluate the unaccounted components of partially accounted findings.

Each of the mechanisms, **aggregation, elaboration** and **projection** are used in the initial formulation of the PSM. Focal aggregation and elaboration create mappings

between nodes across different levels, and causal aggregation and elaboration create mappings between causal links across different levels.

The knowledge representation formalism and operations described above are considered to be sufficient for dealing with effects with multiple causes and feedback loops common in the physiological regulation of the body's vital functions. The mechanism developed is intended for symbolic description for reasoning with and explaining the abnormalities in physiological regulation in a patient, not for predicting the behaviour of physiological parameters over time using dynamic simulation techniques.

3.1.4 DIAGNOSTIC CLOSURE (DC)

The diagnostic closure provides the program with an ability to evaluate the consistency of a finding before it decides to accept it. If the incoming information is true, a major re-analysis of the understanding will have to be undertaken. Therefore, the program has an opportunity to suspend the global diagnostic processing and revert to local processing to validate the finding or to justify ignoring it.

Diagnostic planning generally begins with the global task of discriminating between the alternate explanations provided by the set of PSMs. This task is decomposed into smaller tasks using the following diagnostic strategies:

confirm, differentiate, rule-out, group-and-differentiate and explore.

The diagnostic algorithm for the ABEL is:

1. Presenting Complaints: The serum analysis and the initial complaints are analysed. A small set of initial PSMs is created and added to the list of causal hypotheses (the CH-list).

2. Rank Ordering Hypotheses: All PSMs in the CH-list are scored for quality of explanation they provide for the patient's illness. The leading one or two of these PSMs are selected as possible explanations.

3. Computing Diagnostic Closure (DC): DCs for the selected PSMs are computed and disease hypotheses in each DC are scored.

4. Termination: If the DCs for all PSMs are null or if some PSM provides a complete and coherent account for the patient's illness then the current phase of diagnosis is complete.

5. Diagnostic Information Gathering: Based on the number of DCs, a top level *confirm* or *differentiate* goal is formulated. Using diagnostic strategies, this goal is decomposed

into simpler sub-problems until individual questions are formulated.

6. Restructuring the PSM: If (5) results in new finding, this finding is incorporated into each of the PSMs by extending the structure of the PSMs to take the observed finding into account. This process is repeated starting at (2).

3.1.4.1 Scoring the PSM

The score of a PSM measures the degree of incompleteness of the PSM as an explanation of the patient's illness. It is computed by summing the severities of partially and fully unaccounted states in the PSM. The *Patil* suggests that this algorithm could be improved by:

a) taking into consideration the need of a finding to be accounted for by an acceptable diagnosis;

b) by taking into account the degree of explainability of a PSM.

3.1.4.2 Scoring a Disease Hypothesis

First, they are grouped according to the number of unaccounted findings that can be accounted for by each hypothesis. Second, among those hypotheses that can be accounted for the same number of findings, the diseases

are rank-ordered by a score computed from three factors, which are:

1. match; the number of causes and findings in the PSM that are consistent with the disease hypotheses,

2. mismatch; the number of causes and findings in the PSM that are inconsistent with the disease hypotheses, and

3. unknown; the number of unobserved findings predicted by the hypothesis which are not inconsistent with the PSM.

A disease hypothesis is eliminated from immediate consideration (for one cycle of diagnostic inquiry) if the difference of match and mismatch is below an arbitrary threshold. The match combined with the unknown corresponds to the maximum possible score attainable by a given disease hypothesis. If this score goes below a threshold, the hypothesis cannot be confirmed even if all the remaining unknown findings are resolved in favour of the hypothesis.

This criterion is purely structural. The author believes that incorporation of probabilities as a secondary scoring criterion would substantially improve the quality of the scoring mechanism.

To sum up, a diagnostic closure is created by projecting appropriate states in the PSM or hypothesised diseases forward to identify their predicted consequences and backwards to identify their possible causes.

The information gathering process of each diagnostic cycle is followed by the revision of the structure of each PSM, making it consistent with the newly available information.

3.1.5 LIMITATIONS

The inherent size and complexity of the domain has forced the developers to limit the scope of the research to just a few issues. Even within this limited scope there are some major problems.

The representation of the relation between states is inadequate; all interactions are described using a single type of link, i.e. causal. This is unnatural when there is no known causal explanation. Furthermore, there is a need to group states which jointly have significant diagnostic and prognostic implications even if the states are not causally or statistically related. For instance, associational links and grouping links are needed to capture these cases.

The program also fails to ascertain the overall state of the patient's health, e.g. vital signs, stability etc.

3.2 USE OF AI AND MATHEMATICAL RELATIONS AI/MM

This is the work of Kunz (1984), on integrating simple mathematics and Artificial Intelligence techniques to develop and analyse a physiological model of the renal system.

The program analyses physiological behaviour and explains its analysis. It considers relevant data, identifies whether the data are abnormal and predicts possible effects of any abnormalities.

3.2.1 INTRODUCTION

The physiological model is based on anatomical knowledge, the behaviour of the physiological system and the mechanisms of action of the system.

Knowledge of physical laws is represented mathematically and included in the knowledge-base. The knowledge-base also includes knowledge of anatomy, physiological functions, and measurable parameters of physiological function. The knowledge-base also includes inference rules which are based on a definition of the causal relation between events.

The program uses this knowledge (with the exception of the mathematical knowledge) to make inferences about normal physiological behaviour and the causes and effects of abnormal physiological behaviour. Mathematical

relations are applied if found appropriate and the relation is evaluated either qualitatively or quantitatively as appropriate.

AI/MM makes inferences from knowledge of structure and function. It utilises definitions of causality and heuristic and mathematical descriptions of function.

To define a physiological model, a vocabulary is used that describes processes, substances, parameters, mechanisms of action, underlying bases for describing mechanisms as well as anatomy (structure).

Causality is represented explicitly by rules in terms of events. The "bases" for these causal relations are their underlying principle and are used to provide explanation of their use. The bases include widely accepted empirical observations and laws of physics. There are, therefore, two kinds of causal relations described as "Type-1" and "Type-2" respectively. Type-1 bases for causal relations may have qualitative or quantitative forms; whereas Type-2 can only have quantitative form.

Causal relations may be propagated through an anatomical network to cause a series of resultant physiological events.

There is a hierarchy of bases, mechanisms and processes. This hierarchy provides a strong focus of attention of heuristics for analysing a physiological model to describe and predict behaviour. Problems are looked at in terms of this hierarchy. The problem solution then is to match the abstract patterns of mechanisms with the data for each case. This hierarchy also assists in the process of knowledge acquisition about a problem.

AI/MM, therefore, uses both symbolic knowledge and mathematical knowledge. Mathematical knowledge is used to clarify ambiguities when no well-defined symbolic knowledge is available and to impose constraints on the behaviour of the overall system. These constraints are based on the laws of physics. In addition, certain parameters can be estimated using this kind of knowledge where a direct measurement is not possible.

AI/MM performs the following functions by request:

1. Report the value of a parameter (both qualitative and quantitative);

2. Identify methods for measuring the quantitative value of a parameter or a related set of parameters. These are either stored or inferred from principles of physiology;

3. Obtain a qualitative or quantitative value for a parameter;

4. Interprete the significance of a parameter with a specified value. Predict the effects of an abnormality and therapeutic goals;

5. Print the definition of a concept.

MI/MM reasons about physiological behaviour, identifies abnormality, identifies possible therapies and predicts the potential outcome of therapy.

3.2.2 KNOWLEDGE REPRESENTATION

A knowledge representation system (MRS) Genesereth et al. (1980), is able to retrieve facts from knowledge and databases, interpret rules, and store data and conclusions in the database.

The knowledge-base includes three kinds of physiological knowledge:

a) Principles of physiology,

b) Facts and relations, including facts about real objects, parameters, physiological processes, anatomical relations and physiological mechanisms of action,

c) Rules for inferring conclusions about the patient state.

AI/MM represents knowledge of facts of anatomy and physiology as "concepts". Concepts are very similar to frames in the way knowledge is represented. A concept is defined by its name, type, relation with other concepts or entities, parameters etc. There are 125 concepts represented, each having between 5 to 65 features.

Concepts represent two kinds of physical objects: anatomical and physiological substances. Additional concepts specify features of each parameter. Furthermore, concepts define physiological processes, or the rules that can change parameter values. Concepts also describe mechanisms and bases.

AI/MM has rules that define relations between parameters and physiological concepts. These rules can be used to infer relations and values not explicitly represented in the knowledge-base.

3.2.3 META KNOWLEDGE

AI/MM has two forms of meta knowledge:

a) Twenty five rules which identify that some relations have single values, for instance, a parameter has a single qualitative state or a process has a single mechanism, and

b) Seventy five rules to specify the number of possible instances of a given relation that is invariant.

AI/MM uses this knowledge to improve the efficiency of its search. This is the knowledge that is not described explicitly in physiology, but it is implicit in the descriptions of anatomy and physiology.

3.2.4 REASONING

A causal analysis of the effects of some change can be instantiated by the user. The user can, therefore, ask the system to interprete the effect of some perturbation. AI/MM reasons forward from observed cause to hypothesised effect. The system then searches for further effects of the newly hypothesised cause. Propagation continues until no further effects are found or if a feedback loop is recognised. An event, therefore, can be the cause of an effect and so forth.

The system provides a top-level summary of causes and effects at the highest appropriate anatomical level.

A causal relation is plausible if it is known to be logically possible and if an anatomical link exists. The causal relations used are identified to be of Type-1 or Type-2 bases, and the knowledge-base is used to check for abnormalities of parameters related to the relevant type.

The effects of the causal propagation must be consistent with the laws of physics and physiology.

To interprete the effects of an abnormal event, the system searches for primary effects. If any is found, a record is made into the patient-specific database. Secondary effects are then found. The propagation stops when there are no more effects or when a negative feedback loop is detected.

AI/MM reasons at various levels of detail but it displays the top-level reasoning. Other levels may be displayed if explanation is asked for.

3.2.5 CONCLUSIONS

AI/MM is an interesting example of using empirical knowledge in conjunction with well understood principles of physiology and anatomy to construct a model. The empirical knowledge is of two forms: mathematical laws of physics and causal heuristics. The former imposes constraints so that the behaviour of the system is consistent with laws of physics; or in other words with common sense as well as the more sophisticated laws of physics that are not explicitly represented. Causal heuristics are used to cope with uncertainties due to lack of a full understanding of the renal system.

AI/MM is a hybrid model of the renal system which is able to choose an appropriate qualitative or quantitative analytical technique for a particular problem.

Its problem-solving process involves explicit rule-based reasoning, where rules perform a search to find a path through which physiological function can propagate through an anatomical network. In contrast ABEL, uses relatively complicated operators to aggregate nodes in a causal network. This is due to the lack of explicit representation of anatomy and its relation to physiology in ABEL.

AI/MM is developed such that clinical problems can be analysed much in the same manner as a traditional mathematical model, that is, given a cause what will be the effect(s). The challenge, however, is to use this kind of augmented knowledge representation to do the reverse, i.e. given a set of data (effects) what are the potential causes (differential diagnosis).

3.3 LONG'S WORK ON USE OF A PHYSIOLOGICAL MODEL

One of the most impressive works that is currently being developed is carried out as a joint effort by the MIT Computer Science Laboratory and Tufts New-England Medical Centre in Cambridge and Boston MA. The team is headed by Dr William Long, and although they have been very conservative in publishing their work (three papers in six years, Long et al., 1982; Long et al., 1984; Long et al., 1986) what exists is of a high calibre.

Their work is basically built upon the experience of ABEL but restricted to a smaller and more manageable area of cardiovascular disorders and specifically those of heart failure. It is important to note that this work is experimental and still under development.

3.3.1 SYSTEM OVERVIEW

The approach makes use of a causal physiological model for relating clinical and laboratory data to the mechanisms responsible for the patient's disorder and provides methods to aid the user in reasoning from that model about diagnostic and therapeutic questions. The model and methods are accessible, abling one to use the program as a *reasoning blackboard* (or more recently *reasoning network*) to examine the implications of hypotheses and possible therapies.

The work has been directed toward the diagnosis and management of heart failure, where a thorough understanding of the haemodynamic and physiological relations may provide many clues needed to give a proper interpretation of patient data.

The approach uses a partially constrained physiological model to represent the state of the patient. The user has access to this model through procedures to enter data, to allow diagnostic and therapeutic reasoning, and obtain an explanation of the patient state.

The ultimate goal is to develop a program where the user can enter what is known about a given patient, review the implications in terms of what must or must not be true of the physiological state, consider the implications of hypotheses accounting for that state, look for strategies for gathering appropriate clarifying data, and consider possible therapies. This may be characterized as a **reasoning blackboard** for thinking about a patient with cardiovascular disease. The intent is to develop a program to assist physicians to reason about the diagnosis and management of patients with severe or complex heart disease of any aetiology.

The model represents information about the nature of aetiologies, causal relationships, therapies and measurements.

The program is organized into five modules; the physiological model (the central data repository), an input module, a diagnostic module, a therapy module, and an explanation module (figure 3.3).

The input module receives the data about the patient and sets the qualitative parameters in the physiological model. The other modules operate from the physiological model to assess the completeness of the diagnosis and plan ways of improving it; to search for possible therapeutic measures and to anticipate their possible effects, and to explain the model to the user.

The physiological model is central to the program structure. It is a network of nodes representing qualitative values of physiological parameters. For example there are nodes representing *angina* and another representing *high heart rate*. All nodes initially have the value unknown, reflecting the initial state of knowledge about the patient. As the input module acquires and assesses data, some nodes are assigned a truth value, either when the program decides there is sufficient evidence to justify this or when the user decides the value is appropriate. The nodes are connected by a network representing the minimal logical constraints that must exist among them. Logical relations are automatically maintained by a *Truth Maintenance System* (*TMS*) (for a detailed account of TMS refer to McAllester,

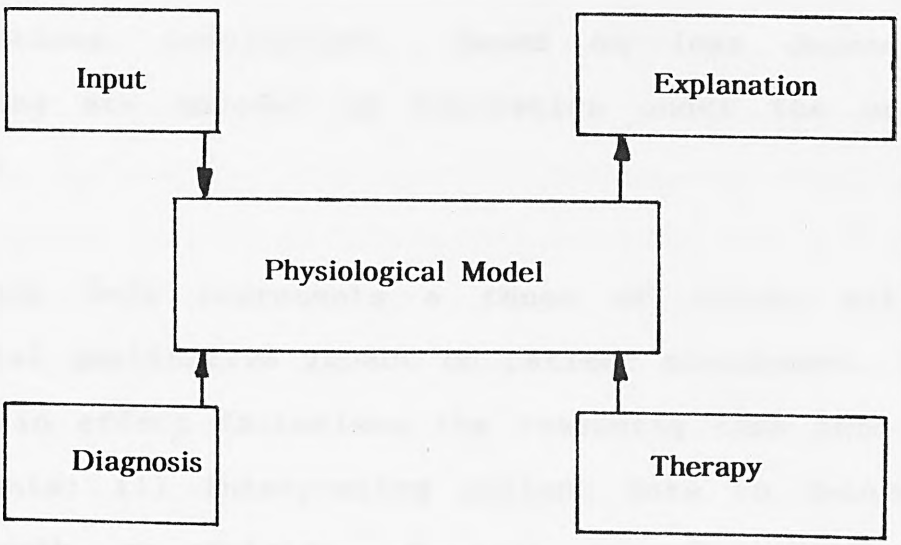


Figure 3.3- Program organisation.

1980) by propagating the implications of each assertion. When an assertion is not consistent with the current state of the model, the inconsistencies are presented so that the user can withdraw any that are not appropriate resulting in the reconfiguration of the logical implications. Conclusions based on less dependable reasoning are encoded as heuristics under the user's control.

Each node represents a range of values with a potential qualitative impact on patient management. This design in effect factorises the reasoning task into two components: (1) interpreting patient data to determine the truth or falsity of various nodes, and (2) determining what diseases and therapies are consistent with the known nodes. Quantitative information is also used to guide the program in gathering input, making diagnoses, and recommending therapy. The nodes represent physiological parameters, primary causes, and therapies. A disease is thus represented by a chain of abnormal nodes implied by various observations about the patient. This kind of representation can handle multiple diseases and multiple presentations of a disease as additions and variations to these chains of nodes. The intent is to distinguish disease states that would change the therapy needs. In addition to the links with possible causes and effects, nodes include links to the possible therapy nodes hence focusing diagnosis and therapy, and a list of possible measurements that might provide evidence for

the truth of the node. Both the therapy nodes and measurements have risks, benefits and requirements that translate into costs to aid in selection.

The diagnostic module attempts to relate clinical signs and symptoms (pulmonary oedema, fatigue, poor renal function, angina, etc.) to both their original cause and possible aggravating factors. The therapeutic module uses these causal chains to identify therapies that may break the chains. Since therapies may have multiple effects or a given effect may have multiple implications in the overall system, the therapies may have effects that will aggravate the patient state. The program analyses the potential effects to assist the physician in anticipating both the expected and the unexpected outcomes.

Long et al. (1986), report the result of development of an algorithm based on the signal flow analysis for predicting effects and implementation for handling multiple effects, changes over time, non-linear relationships and providing explanations. When applied to the physiological model of the cardiovascular system, this methodology predicts drug effects consistent with the medical literature.

3.3.2 PREDICTING CHANGES

The approach chosen is different to either quantitative or qualitative (e.g. Kuipers, 1986) simulation. In the quantitative simulation there is a

need to know the current value of each parameter, many of which are not readily measurable. In the case of the qualitative simulation on the other hand, in the cardiovascular domain there is a tendency for explosion of possible model states when adding opposing influences of unknown magnitude.

The approach adopted by Long et al. is to assume the system will reach a stable state after a perturbation. The question then becomes how that stable state is changed from the state before therapy. Signal flow analysis is used in the domain of circuit analysis to predict circuit gain. In a similar way, after making simplifying assumptions and modifications to the signal flow analysis machinery, one is able to apply this approach to reasoning about a physiological model. The assumptions are as follows:

1. The system goes from steady state to steady state. It is assumed that for the time period of concern, parts with shorter time constants have reached a stable state and parts with longer time constants have no effect.

2. The system can be modelled as being piece-wise linear.

Given these simplifications, the network of physiological parameters is linear and the techniques of

signal flow analysis become applicable. A new formulation of Mason's General Gain formula (Mason, 1956) is derived that computes the gain incrementally from parameter to parameter, correcting for feedback each time a new feedback path is encountered (the derivation and implementation of this formula are to be discussed in a forthcoming paper). Essentially, the computation involves computing for each path from the changed parameters to other parameters the gain for each link. The gain is the inherent gain of the link adjusted for any new feedback loops encountered by the path at that point and the change in a parameter is the sum of the path gains going through the parameter.

The relationships on the links between parameters are formulae determining the link strength from the parameter states that influence it. If the relationship is linear, this is just the strength of the link. The link strengths in the current model can be zero (decoupled) or positive or negative with value either 0.5, 1.0 or 1.5 (the algorithm can support any values). The parameter values are scaled such that these link strengths can be presented as weak, moderate or strong relationships. These values have proved to be sufficient to represent the experimental knowledge of the relations. They are also sufficient to account for the behaviour of the system discussed below.

3.3.3 APPLYING CHANGE ANALYSIS TO THE MODEL

The signal flow analysis algorithm is implemented in stages. All feedback loops are computed upon loading the model; this allows the path generation to determine the loops encountered by the path at each new parameter. The gain along a path is the product of the gains across the links and the total gain at any parameter is then the sum of path gains to the parameter. Multiple changes to the system are handled by summing the changes. The loops and parameters are both represented as bit vectors to increase the efficiency of the many membership comparison operations.

As parameter changes are determined by summing the changes along the various pathways, the contribution of pathways can be compared. To explain a parameter change, the program examines the contributing pathways and highlights the pathway making the largest contribution and those making some threshold as much.

Since the changes caused by therapies are not small, a piece-wise linear approximation must be considered. The program determines which link gains will alter first from the changes and the current parameter values. These are then changed to the next region and changes are recomputed. Hence the total response can be computed. This is not a guaranteed method to make appropriate changes as the transient behaviour may be different from

the stable state, but in a highly damped system such as the cardiovascular system the assumption is reasonable.

Another problem is that different parts of the cardiovascular system take different times to stabilise. The program assumes the pathways with long time constants have no effect on short-term solutions and changes are determined separately for the different time periods. Thus, the algorithm for determining the changes following large dosages of drugs over a long period of time: starts with the shortest time constants in the system, then determines the changes for a small dosage. The parameter values affecting non-linear gains are then changed and recomputed as necessary to determine the immediate changes for the appropriate size dosage. Move to the next time period with the projected parameter values and compute the changes including links that have effects within that time period. Continue until the desired predictions are determined.

3.3.4 APPLICATION OF THE MODEL TO THERAPY

To validate the approach, the predictions of the model in the normal state were compared with information in the literature on the effects of the major classes of drugs used for the treatment of patients with heart failure and coronary heart disease. The drugs are represented by adding each node as a model parameter affecting those parameters *directly* affected by the drug. Some drugs affect a single parameter, others affect a

number of parameters. The changes predicted are represented in direction and relative amount. The predictions are normalised so that the largest change for each drug is represented as three arrows and smaller changes as lesser number of arrows. The model predictions were mostly consistent with the literature.

3.3.5 CONCLUSIONS

The M.I.T./Tufts group have clearly recognised the need for a detailed and explicit representation of knowledge. This knowledge as previously mentioned, is represented in a causal network in effect producing a model of the domain. This model is by no means exhaustive and is being actively updated to cater for the new requirements. However, the beauty of the system lies in the fact that this physiological model is used for all aspects of reasoning, from diagnosis to therapy planning and explanation. This very much resembles the way a clinician reasons when the data is considered in the context of the physiological domain (i.e. conceptual model) and following interpretion appropriate measures are taken. The clinician anticipates some change and is concerned not with the absolute change of the relevant parameters but rather with the relative change. Such a conceptual model helps to explain the observed data (before or after intervention).

One other important aspect of this program is that it is capable of non-monotonic reasoning, since the

program puts every new piece of information into context and re-computes its effects.

For the purpose of efficiency, an algorithm was developed that simplifies reasoning to some extent. Such simplifications are needed to produce systems with acceptable response time (currently a few seconds on a Symbolics 3640). Although the system has not been evaluated with reference to real data yet, it has shown good agreement with the medical literature.

Although the system is an experimental one, the developers have shown a very balanced approach to using qualitative and quantitative knowledge. This will be further discussed in chapter 6.

3.4 A SEMI-QUANTITATIVE SIMULATION FOR REPRESENTATION OF DYNAMIC CAUSAL KNOWLEDGE

The work of Widman is another example of combining quantitative and qualitative techniques and employs simulation to overcome some of the problems of each of these techniques. The approach is a symbolic extension of the system dynamic method which manipulates symbolic descriptions of dynamic systems to predict semi-quantitatively their future states. The work is nicely described in a paper presented at MEDINFO 86 (Widman 1986). The following is a summary of the paper describing the work.

3.4.1 CONCEPTUAL ISSUES

3.4.1.1 Qualitative Issues

The basic assumption is that the causal network contains implicit functional-structural information for which the program must be able to make reasonable self-consistent default assumptions.

In this program, all quantities are defined relative to their own "normal" value. Hence, they can be mapped on to qualitative adjectives such as "high", "very low" etc. Unlike in qualiitative methods where the set of permitted values is closed under the arithmetic operations, the set of allowed values here is not bounded

by the method. Thus "large" plus "large" is not equal to "large", but two times "large".

The adjectives, when defined, are mapped by the program on to the real numbers in the interval (-1,..,0,..,+1), where 0 is normal, -1 is 100% below and +1 is 100% above normal. To deal with the problem of ranges, the program uses the mean value for each adjective.

3.4.1.2 Numerical Issues

In order to deal with the problems arising from error due to truncation and any increase in the size of the time interval due to the polynomial approximation used to perform numerical integration, some of the work of Guyton et al. (1984) is being used. Specifically these are:

a) segmentation of the model; with iteration to steady-state using time intervals proportional to the time constants of the segments, so that short-time constant loops are brought to steady state and the longer iteration time intervals are used for the long-term loops.

b) automatic shifting of the time frame based on the degree of oscillation in each integrated variable (adaptive integration).

3.4.1.3 Modelling Issues

The explicit information is interpreted and the model instantiated. The variable types which the program knows about are defined (see below).

3.4.2 SYSTEM OVERVIEW

3.4.2.1 Definitions of the Causal Network

(a) Material versus Information.

All network variables are either "material" or "informational". "Material" variables are conserved quantities; "informational" variables pertain to regulatory mechanisms or to quantities which need not be conserved.

(b) Network Building Blocks

All network variables are either "material" or "informational". "Material" variables are conserved quantities, such as mass, momentum or energy. "Informational" variables pertain to regulatory mechanisms.

(c) Properties of Relationships or Linkages

Relationships or linkages between variables represent processes. The only required properties are the direction of linkage and arithmetic sign. Additionally, temporal relationships, linear quantitative relations

and special parameters for specific variable types may be specified.

3.4.2.2 Specification of Initial Conditions

The initial values of all variables in the network must be available for the simulation to proceed. Currently, if the values are not known, they are considered to be normal. A better approach however, is to make logical inference of the available data and the known causal relationships to form "coherent hypotheses" as proposed by Patil (1981), each of which includes all known values and postulates a self-consistant set of values for all other variables in the network. This method is currently being implemented.

3.4.2.3 Formulation of Equations

The network specification is translated into difference equations automatically using the definitions of the variable types.

The Euler difference equation approximation is used.

3.4.3 IMPLEMENTATION

The program is currently written in MACLISP and has been translated into VAX COMMON LISP. The program requires 5 to 10 minutes to run 20 iterations of a 34 variable network on a PDP-10.

3.4.4 RESULTS

To test the accuracy of knowledge representation techniques, an application area in a medical domain was chosen. The domain is that of cardiovascular sytems. This is sufficiently rich in levels of detail and there is also vast experience with the plausible perturbations of the domain.

A network consisting of symbolic representations was built along the lines of well-tested numerical models of the cardiovascular system. These programs obtain their results by numerical integration of differential equations, as is performed in this method. However, their models are not symbolic. The current causal network has extra detail only at variables whose patterns of behaviour depend strongly on relative quantities. It is at these variables that qualitative methods face the combinatorial problem and the problem of generating behaviour which does not actually occur. One of the strengths of this method is that extra information can guide the simulation.

The performance of the current causal network was tested on sixteen classic disorders, each of which is known to be accounted for entirely by the disturbance of one or at most just a few initial values. The same network was used for all simulations. Simulation of all of the examples yielded acceptable results by the

criterion of semi-quantitative agreement with the medical literature.

3.4.5 CONCLUSIONS

This work is a good example of a balanced approach in combining qualitative description with a quantitative simulation to model a domain. Both qualitative and quantitative simulations alone have their shortcomings.

A symbolic representation is desirable because of its explicit and detailed representation, however, after a period of simulation from a steady state to another there is no information about the intermediate stages. Moreover, representation of a complex system such as cardiovascular system in a purely qualitative fashion may give rise to unrealistic predictions because of the large number of feedback loops present, where adding opposing influences of unknown magnitude is a problem. In other words, direction of change is not sufficient and some information regarding the magnitude is needed.

The system therefore, starts by symbolic representation of the domain, using generic functional "building blocks" or variable classes, and predefined default values (where there is no information available); translation of the model into a system of first order difference equations and integration of the equations. The output can be translated back to symbolic form by feature extraction, but this is not yet implemented.

3.5 MODEL

This is an experimental work that is being carried out at the Royal Free Hospital School of Medicine. The objective is to develop a software package that permits development of a model of any domain (that can be represented adequately by compartments) from the conceptual description. This description is transformed into symbolic equations and a general-purpose algorithm written in Pascal is used for the numerical simulation of the symbolic equations.

MODEL is not a decision-making tool or a knowledge-based system, rather it is a tool for development of physiological models (qualitative and quantitative). It is reviewed because of its novelty and its potential for development of a knowledge-base from the conceptual description of the domain. As mentioned in Chapter 1, clinicians have a conceptual model of their domain of expertise. Hence, the knowledge acquisition process can be speeded if a tool is available to translate a conceptual model into a computer model (be it qualitative, quantitative or a mixed approach).

The work is described in a recent paper (Leaning and Nicolosi, 1986). As mentioned above, MODEL has the capability of manipulating conceptual descriptions and symbolic mathematical equations to perform conventional

numerical simulation. Such representation is defined as a 'knowledge-based model' by the authors. In MODEL:

a) facilities are provided for the user to build a conceptual or linguistic description of a compartmental system in a physiological domain. This description consists of the anatomical site, type, amount of substance and flux (the exchange of substance between compartments);

b) a set of symbolic differential equations is generated automatically from the conceptual description which forms a qualitative constraint network;

c) a modified version of the QSIM algorithm (Kuipers, 1985) is used to perform a qualitative simulation of the constraint network;

d) numerical simulation can be carried out using the symbolic differential equations, provided the user supplies the appropriate initial numerical values.

3.5.1 A MULTILEVEL SCHEMA

Based on the previous experience of one of the authors in modelling methodology (Leaning, 1980), a framework for knowledge-based modelling was devised.

This methodology starts with the problem perception and model specification to arrive at the model formulation and consequently at identification and parameter estimation. This process leads to a fully quantitated model. A backward step may be taken at any point to modify any of the steps.

Model formulation consists of conceptual description and hence derivation of differential equations and the relevant simulation. Conceptualization identifies the key elements of the system and their interaction, both structural and functional. A natural form of representation is associated with each step of modelling and its related information with a distinct level.

3.5.1.1 Realisation of the Scheme

The built-in control is based upon the idea that the user builds a model at the conceptual level and then proceeds to symbolic and numerical levels, providing necessary further information when required.

MODEL is written in LPA PROLOG with the simulation engine written in Pascal. It runs on an IBM AT running MS-DOS. User-interaction is through a menu-driven system with associated nested windows.

3.5.1.2 Specification and Control

Model specification consists of the description of the system to be modelled and the level of detail and

accuracy required. The purpose of the model is also stated. Therefore a quantitative or qualitative simulation will be produced based on the purpose for which the model is being developed. In this way the program controls the depth and detail of the representation.

3.5.2 THE CONCEPTUAL LEVEL

The components and structure of the compartmental system are defined at this level, with details of physiological sites and substances. The model need not be physiologically complete before exploring the symbolic or numerical levels.

3.5.2.1 Internal Representation

The conceptual description is represented as a set of Prolog predicates. There are seven basic forms: compartment; flux; loss; input; modulator; measurement; order. An index is generated internally and is associated with each of the above forms indicating the relevant compartments. For a detailed description of the forms see Leaning and Nicolosi (1986).

3.5.3 THE SYMBOLIC LEVEL

At this level the compartmental system is symbolically represented in mathematical terms. Differential equations are thus represented symbolically from the conceptual description. Qualitative or

quantitative simulation is then carried out depending on the purpose of the model.

The differential equations for the compartments are represented as Prolog predicates using an index number to identify the various compartments.

3.5.3.1 The Qualitative Constraint Network

The qualitative constraint network is based on the QSIM algorithm (Kuipers, 1985). QSIM simulation is represented as a sequence of qualitative states. Of course, there is no information available between the time intervals. Each variable is treated as a function mapped from a finite ordered set or "quantity space". The entries in the quantity space are known as land-mark values. Direction of change of a function must be given; two states are distinct if their qualitative values or direction of change or both are different.

The generated states are checked for consistency by re-expressing the model in a constraint network.

MODEL automatically generates the constraint network and its initial state from the conceptual description.

A problem associated with qualitative simulation is that of "branching". That is, the algorithm may arrive at unrealistic possible states as well as the realistic

one(s); additional information can be inserted to avoid such branching.

MODEL is capable of simulation with incomplete knowledge.

3.5.4 NUMERICAL LEVEL

The symbolic differential equations are interpreted and the essential initial values and parameters asked for. A general purpose simulation program written in Pascal, (Leaning, 1986), is then used to perform simulations. The results are displayed in graphical form for the compartments that have been defined at the conceptual level (as a measurement term).

3.5.5 CONCLUSIONS

This work is the result of the experience and the awareness of the developers of the problems associated with numerical simulation both at a conceptual level and at the implementation level. The objective therefore was to devise a scheme by which some of the problems in development of simulations could be overcome.

MODEL is one of the rare (if not unique) environments where one can develop a model from conceptual description and run qualitative and quantitative simulation. This provides the facility to develop alternative models and hence an optimised model could be achieved very quickly.

Currently, the simulations are displayed graphically and no attempt is made to assess the results in a qualitative manner; but research is being carried out to examine different methods of assessment.

Another way of enhancing the system could be by looking at ways to achieve some interaction between the qualitative and quantitative simulations. As the MODEL stands, the simulations are quite independent of each other.

In all, the approach is a very exciting one that could change the realm of modelling not forgetting its potential as a tool to develop a knowledge-base from a conceptual description of the domain.

3.6 KARDIO

KARDIO (Bratko et al., 1986) is a system that interprets ECG signals for cases of cardiac arrhythmias; it is also capable of performing the reverse procedure of predicting possible ECG signals for cases of arrhythmias.

The system is based on a qualitative model capable of simulating various cardiac disorders either singly or in conjunction with others. This system is particularly interesting because:

1. it highlights the capability of qualitative models to deal with multiple disorders, and

2. a practical approach is taken to produce an efficient system without losing much detailed knowledge.

Knowledge in KARDIO is represented symbolically and reasoning is carried out from first principles, thus yielding a qualitative model. The model is then simulated to arrive at various diagnostic and predictive states which could be used for the two purposes. But for the purpose of efficiency inductive learning programs were used to compress this knowledge-base.

Each of the above constituents are discussed below.

3.6.1 THE QUALITATIVE ("DEEP") MODEL

This is a qualitative model of the electrical activity of the heart which is represented as causal relationships between objects and events in the heart. The model is analogous to an electrical network but the signals are represented qualitatively by symbolic description.

The model consists of:

1. Nodes, which are comprise; impulse generators, conduction pathways, impulse summators and ECG generators; all of which are represented symbolically.

2. A dictionary of simple arrhythmias related to heart disorders. These are defined in terms of functional states.

3. "Legality" constraints. These are states that are rejected by the model corresponding to the following criteria:

    a. Logically impossible states
    b. physiologically impossible states
    c. medically uninteresting states.

    In this way the search space is considerably reduced.

4. "Local" rule sets. Specifying the behaviour of the individual components of the heart in the presence of abnormalities.

5. "Global" rules. These are rules defining the causal relations between various components of the model and ECG features, thus reflecting the structure of the network. There are 35 global rules in the model represented in PROLOG clauses. A detailed description of the model can be found in Mozetic et al., (1984).

3.6.2 THE QUALITATIVE SIMULATION ALGORITHM

In order to insert additional control during the execution of the program, an algorithm is used instead of the PROLOG interpretion mechanism.

The simulation algorithm allows theorem proving and theorem generation.

A simulation run consists of:

1. Instantiation of the model by a given arrhythmia.

2. Checking the functional state against the legality constraints.

3. Execution of the model by triggering rules until no more rules are fired.

4. Collection of the proved assertions about ECG signals followed by an ECG description corresponding to the given arrhythmia.

The most natural way of implementing the above algorithm is to use the depth-first search strategy. Although this is efficient for prediction type queries, diagnostic type queries run into trouble, since the model is running backwards, and branching becomes a major problem.

Various methods were tried to compensate for this inefficiency by first rewriting the model and introducing more constraints to limit the branching. This, however, increased the size of the model considerably and also affected its transparency which is of great importance for the explanation of its behaviour. Another alternative was to generate ECGs for all possible arrhythmia cases and store these as associations. This is also inefficient because, for each disjunctive solution the simulator has to backtrack to some previously used rule in the model and restore its previous state; also, the resulting ECG descriptions have the form of disjunctions of ECG patterns, which can be more complex than necessary and can be simplified later. But, the simplification procedure is again a

complex operation. The simplification can be carried out when a disjunction is generated and before it is further expanded; this is a more economical simplification.

Because of the two factors mentioned above, another implementation was sought that could handle alternative execution paths in a breadth-first manner. The result was an algorithm that would generate parallel alternatives and simplify disjunctions. The simplification rules, however, are local, i.e. model dependent, therefore the "breadth-first" simulation is not general and the simplification rules need to be modified if there is a change in the model.

Bratko reports that this specialized simplification proved to be quite powerful compared to the depth-first simulation. The depth-first simulation generated 72 ECG descriptions for the combined **arrhythmia atrial fibrillation** and **ventricular ectopic beats** as opposed to only 4 generated by the above algorithm.

Using the breadth-first simulation algorithm, the knowledge-base was generated automatically by executing all **mathematically** possible combinations of simple arrhythmias. A large number of the combined arrhythmias were eliminated by the legality constraints over the state of the heart. In this way the knowledge-base (i.e. the arrhythmia-ECG base) is complete in the sense that all possible physiological and mathematical cases are

present. The diagnosis process is now the simple task of finding (retrieving) the arrhythmia(s) that correspond to an ECG description.

The knowledge-base generated, despite being complete, suffered from one main drawback- its size. There were 8314 PROLOG clauses occupying 5.1 Mbytes of memory. In order to achieve a more compact knowledge-base, **inductive learning algorithms** were used. However, the knowledge-base was far too big for these algorithms to learn from examples. Therefore a sub-set of the knowledge-base was derived which was complete in its own right, and the cases discarded would be regenerated by adding a few additional rules.

The learning sub-set was substantially smaller the original one comprising 586 combined arrhythmias and 2405 ECGs compared to 2419 arrhythmias and 140966 ECGs. This subset only takes up 400 kbytes of storage.

The performance of the system using the inductive algorithms corresponds very well to the definitions in the medical literature, although in some cases much more detailed specification is generated, which may not be necessary for a medical user.

3.6.3 CONCLUSIONS

The novelty of this work without doubt lies in the representation of knowledge and the subsequent use of it

to provide an efficient diagnostic system. The developers are consistent and careful throughout the development of various stages, firstly to represent a qualitative model that is sufficiently rich in its physiological representation knowledge; and secondly to generate all the mathematically possible and physiolagically allowable cases automatically; and finally to compress the required knowledge-base without losing detail or any possible cases.

KARDIO was evaluated with a **select** population, and it could handle 75% of the arrhythmia cases correctly. In an actual test on 36 random cases it could handle 34 cases (94%). The cases where it failed were due to some incompleteness of the deep model, such as incapability to handle an artificial pacemaker.

The deep model can be used for other purposes such as providing an explanation.

3.7 SUMMARY

When studying the evolution of expert systems in medicine, an interesting shift of emphasis is observed.

In the early years, systems were developed based on purely associational knowledge (eg. Shortliffe, 1976), where signs and symptoms were related to the underlying disorder(s). The knowledge used for this kind of representation was based on both well understood mechanisms and purely heuristic knowledge. However, they were both represented heuristically. This to a great extent degraded the ability of the system to reason "intelligently" as the reasoning mechanism was essentially that of pattern matching; as well as not being able to produce an acceptable explanation of its behaviour.

Then, there was the emergence of systems whose knowledge were based upon the principles of causality (e.g. Patil, 1981). Experience of such systems showed relating physiology and medical decision making to be a far too complex and ill-understood domain that could be represented in that manner adequately and efficiently.

Recently, we have witnessed the emergence of systems which combine both qualitative and quantitative techniques, as well as using additional techniques. Some of these systems have been reviewed above. What has become apparent is that clinicians use different

knowledge at different levels of detail at different times. For a knowledge-based system to be able to perform as well as an expert, these knowledge sources must be available to it. Having various knowledge sources available and using them appropriately are, of course, two different problems.

The objective of this review using specific examples is show the complexity of some of the medical domains and the way that different workers have approached the problem. There is no simple solution to the problems involved, but it is the view of the author, that combining and utilising these various knowledge sources appropriately is the way forward.

In the next chapter, the aims, objectives and requirements of the work carried out as well as a detailed description of the developed prototype will be presented.

CHAPTER 4

[redacted]

(Francis Bacon)

4.0 SYSTEM DEVELOPMENT

Having reviewed some of the earlier relevant work in the previous chapter, in this chapter the requirements, choice of model, various difficulties encountered, and a detailed description of the structure of the two versions of the developed prototype will be described. A more detailed discussion of the methodology and associated problems will be discussed in the next chapter.

A brief description of the first version of the developed prototype (MK I), can be found in Shamsolmaali et al., (1988), and Shamsolmaali et al., (1987).

4.1 INTRODUCTION

Clinicians use different levels of knowledge at different times according to the complexity of the problem at hand and its nature. The performance of a knowledge-based system should, at least in theory, be enhanced if it has access to additional sources of knowledge. The objective of the exercise described here, was to explore the extent to which a mathematical model could enhance the performance of a knowledge-based system in a clinical environment.

4.2 REQUIREMENTS

The requirements are based on the original proposal to the DHSS for the funding of this work. A prototype was required that could investigate the possibility of utilizing a mathematical model in a clinical setting, and to explore the advantages of such a system. It was considered that the prototype should be capable of offering "opinions" to the user about the assessment of the current state of the patient and to suggest appropriate treatment justifying it using the model. It was proposed that an "off-the-shelf" model would be used, rather than a "tailor-made" one, as considerable work had already being carried out within the group on the modelling of the relevant domain (Flood, Carson and Cramp, 1985).

4.3 CHOICE OF THE MODEL

As mentioned above, some work had been carried out in the group on the modelling of fluid and electrolyte metabolism. However, after close examination of the implemented model, it was concluded that further work was required to enhance its usability and, more importantly, to validate the model. It was considered important to utilise an established and validated model (if possible) to meet these requirements.

There were two other models available to the group HUMAN (Guyton et al., 1984) and Macpee (Dickinson et al.,

1987). Macpee was chosen because it is well-established (developed about 1972) and although it had never been formally validated it had, however, been in use for a number of years as an educational tool and undergone constant updating. More importantly, as far as the implementation of the prototype was concerned, the developers of Macpee were based in London enjoying a good working relationship with our group. This proved to be of great importance in the following months.

4.4 DESIGN DIFFICULTIES

One of the major drawbacks of Macpee (and indeed many other models) is that it is parameter driven rather than data driven. This limits the process of tuning the model to a particular patient as data cannot be input to the model. The model simulates a cause-effect cycle for a "normal" person. The notion of "normality" is defined as a 29 year old male of 70 kg weight and 175 cm height. An assumption is made that human's circulatory and fluid metabolism behaves similarly under influence of disorders and in health regardless of individual characteristics of patients. It is important to note that as far as this project was concerned, after a simulation, a qualitative change of parameters from what is being perceived as the "steady state" before the simulation, is considered rather than a quantitative number to number correspondence of model parameters with that of the patient.

4.5 SYSTEM STRUCTURE

The basic structure of the prototype is shown in figure 4.1. The system consists of six components which are described below.

4.5.1 Overall Structure

The prototype consists of a user interface, patient database, patient record file, diagnostic module, treatment module, and the dynamic mathematical model Macpee.

Mathematical models can be incorporated within a knowledge-based system in a number of ways. They can constitute an external model with parameter exchange; can be semi-integrated with part of the model used as part of the knowledge-base; or can be fully integrated with the entire model constituting a part or the whole of the knowledge-base (Nicolosi, 1986). If it is to be fully integrated, the model should simulate the pathophysiology of the relevant domain in an adequate manner, and furthermore, it should be able to simulate the state of a patient closely.

The model used in this project, Macpee, cannot be tuned directly to a specific patient since it does not accept laboratory data as input. Therefore, it is used as an external module with interaction between model and knowledge-base being reduced to parameter exchange as described below. Some design features of such a system are outlined in Cramp et al. (1985).

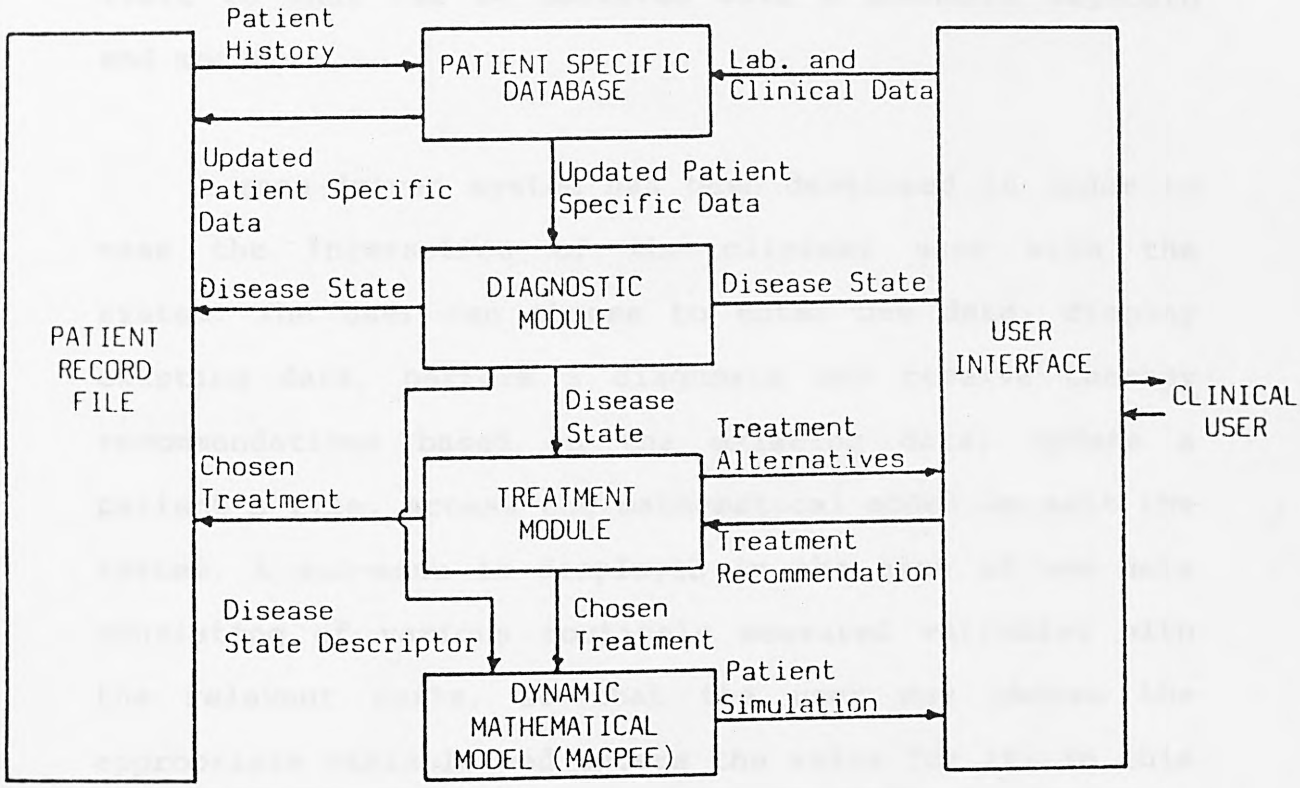


Figure 4.1- Schematic diagram of the overall structure of the prototype.

4.5.2 User Interface

Special attention had to be given to the user interface, as one of the objectives of the project was to develop a prototype that would be used in a clinical setting where it would be further validated. There is limit to what can be achieved with a standard keyboard and monitor.

A menu-driven system has been developed in order to ease the interaction of the clinical user with the system. The user can choose to enter new data, display existing data, perform a diagnosis and receive therapy recommendations based on the existing data, update a patient's file, access the mathematical model or exit the system. A sub-menu is displayed in the case of new data consisting of various routinely measured variables with the relevant units, so that the user may choose the appropriate variable and enters the value for it. In this way typing mistakes are minimised.

4.5.3 Patient Specific Database

The patient specific database is comprised of clinical (bedside) and laboratory data. Current data input to the system by the clinical user is stored and those contained in the record file for the specific patient are retrieved. The variables included are listed in Table 4.1.

Table 4.1- Variables in the Patient Specific Database

Laboratory data:	Plasma Concentration of-	Sodium
		Potassium
		Albumin
		Creatinine
		Urea
	Urine Concentration of-	Sodium
		Albumin
		Potassium
	Plasma Osmolality	
	Urine Osmolality	
	Haemoglobin	
Clinical Data:	Blood Pressure	
	Central Venous Pressure (CVP)	
	Pulmonary Capillary Wedge Pressure	
	Temperature Difference (Core- Periphery)	

Data are entered into the system via the menu-driven mechanism and stored as facts in the patient file. In order to differentiate between data input at different times, an indexing method has been developed. The data have the following PROLOG form:

```
data(Patient-Name,Index,Item,Value).
```

The predicate *data* is used to store the facts about a particular patient. This predicate has four arguments. The variable *Patient-Name* is assigned to the name of the patient (or it could be assigned to the patient's hospital number). The variable *Index* is set to 1 when there is no previous record on the patient, and is incremented by 1 on each subsequent interaction. The most recent data are used for processing.

4.5.4 Patient Record File

Once the name of the patient is entered, the program searches for a file by that name with an extension of ".LOG". If one exists, it is loaded and any new data input will be saved in this file, with the index incremented by one. If there is no file by the name of patient, one is created. At the end of the consultation session the program asks the user whether the new data are required to be saved and acts appropriately. In order to eliminate a case problem with regard to the file name (i.e. lower, upper, or mixed), the typed name is changed

to upper case (using ASCII codes for inversion) before the program attempts to search its record file.

4.5.5 Diagnostic Module

Before any diagnostic action is taken, the numerical data are classifies by comparison with a set of reference ranges. This classification categorises the data into one of five regions namely, low, moderately low, normal, moderately high, and high. The classified data are stored in the RAM memory and displayed, so that the user can follow the steps that are being taken by the program.

When the user has finished inputing data to the system, the most recently classified data (i.e. the ones with the highest *Index*) are collected as a list, called the *characteristic-list*. The *characteristic-list* has the following PROLOG format:

```
characteristic-list(Patient-Name,[[plasma-Na,low],
                                  [plasma-K,normal],
                                  [urea,high]]).
```

The disorders are represented by their name and the possible associated pattern of signs and symptoms having the following PROLOG format:

```
disease-state(water-overload,[[plasma-Na,low],
                              [plasma-Na,low-normal],
                              [plasma-K,normal],
                              [plasma-urea,low],
```

```
                                    [creatinine,low],
                                    [albumin,low-normal],
                                    ........]]).
```

This represents the possible combination of signs and symptoms that are not inconsistent with the disorder *water-overload*.

To produce a diagnosis, the **characteristic-list** is checked against the **disease-states**. The **disease-state(s)** whose pattern of signs and symptoms do not contradict the pattern of signs and symptoms of the patient are displayed as differential diagnoses. There are 15 disease-states defined in the knowledge-base, all of which are water and salt related disorders. The disease-states represented are listed in Table 4.2.

Not all the signs associated with a particular disease-state need be present for it to be a member of the differential diagnoses list. In fact, the reasoning mechanism is that of elimination of disease-states which suffer from a conflict with the characteristic-list. This reasoning mechanism has the advantage of considering all the disease-states without any bias or weighting.

One disadvantage, however, is the fact that the program searches blindly through its knowledge-base to come up with potential disorders. This blind search can be compensated for at the representation level of the

Table 4.2- Disease States Defined by the System

Water overload
Water overload with renal failure
Syndrome of inappropriate antidiuretic hormone release (SIADH)
Congestive cardiac failure
Nephrotic Syndrome
Loss of sodium with water replacement
Loss of sodium with water replacement on a diuretic
Normal plasma sodium with overload
Post-operative normal plasma sodium with water overload
Over-transfusion of blood
Normal plasma sodium with dehydration
Diuretic induced dehydration
Hypernatraemia due to hypotonic fluid loss
Hypernatraemia due to sodium overload
Addison's disease

disease-states without altering the search mechanism. This is done by inserting constraints on the possible combination of signs and symptoms of a particular disorder, so that an unacceptable diagnosis can be avoided. This is an implicit representation of knowledge.

This method of refinement of the knowledge-base appears to be sufficient for conditions arising from a single disorder.

The program is, however, unable to cope with conditions arising from multiple disorders being present. The reason for this failure becomes apparent when a close inspection of the knowledge representation is made. Each disease-state is an entity in its own right without any links or relations with other disease-states. There are no other rules to describe a link or a relation at the reasoning level either. This means conditions arising from a conjunction of disorders present can only be represented explicitly as appropriate disease-states. This problem will be discussed further in the next chapter.

Once a diagnosis is established, it is used to change the parameters of the model, using appropriate rules, so that Macpee is tuned to the diagnosed disorder. The diagnostic module, therefore, acts as the front-end to Macpee as well as assessing the condition of the patient.

4.5.6 Treatment Module

The treatment regimen produced is not based on the outcome of the diagnostic module, contrary to the implication of figure 4.1. The link between diagnostic and treatment modules is left out deliberately, so that the performance of the two modules can be checked against each other as well as reducing the dependency of the overall performance of the prototype on the diagnostic module.

In contrast to the diagnostic module, the knowledge is represented as a set of rules. These rules associate the abnormal data with the appropriate therapeutic measures. The abnormal data are grouped together in the context of the underlying (sub)system. Hence, the treatment module considers various subsystems separately and provides strategies to correct the abnormality.

The sub-systems considered are the cardiovascular and renal systems together with the dietary state and drugs already being taken by the patient. At the end of the inspection of the sub-systems, the program advises on the amount of appropriate intravenous fluid to be given to the patient.

The cardiovascular data (parameters) checked are: diastolic pressure, central venous pressure(CVP) and pulmonary capillary wedge pressure(PCWP); inotropes and

colloid are prescribed accordingly. The renal data checked are: urea, creatinine, their ratio, together with urine sodium and urine output.

Some typical rules are:

IF Diastolic Pressure is Low
and CVP is Low
and Plasma Na in the range 145-150
THEN Prescribe:
Half Dextrose 5% + Half Colloid (volume)

IF Urine Output > 40 ml/h
and patient NOT on diuretics
and Urinary Na > 20
and Creatinine elevated
THEN Renal Impairment

The first rule uses the cardiovascular data and the value of the plasma-Na to suggest the appropriate fluids. The rule also suggests, implicitly, that there is a negative fluid balance. The amount of fluid to be given is decided by other rules based on the state of the fluid balance and the weight of the patient.

The second rule is more interesting; first it excludes the possibility of diuretics to be the cause of the large urine excretion and then checks the value of

creatinine and urinary-Na to establish a conclusion. This is added to the data record of the patient.

The recommendations of the treatment module are passed to the model using appropriate rules (discussed in the next section).

4.5.7 Macpee

The dynamic mathematical model incorporated in the prototype is Macpee, one of the Mac series of interactive digital computer simulation programs designed to assist students to learn about the physiology of major body systems in health and disease.

The output of the model is normally in the form of a graph of blood pressure and the pulse rate plotted against time. At the end of each run, normally 24 hours but this can be changed, variables values are printed out.

The model will perform as many simulated hours or days (maximum 10 days at a simulation cycle) as is directed, and can then be stopped to allow the operator to make changes (within physiological constraints) in any of the large number of variables. In addition, any parenteral fluid normally available in hospital practice may be administered or discontinued (eg. 5% dextrose, saline, potassium chloride solution, blood, packed cells). Similarly oral intake of fluids can be either

restricted or augmented to any specified extent by simulated instructions to nurses.

The adoption of Macpee as the dynamic mathematical model in the prototype design was a compromise. On one hand it is a rich physiological analogue which has been widely used in an educational context. As such it could be incorporated immediately thus circumventing the development cycle of a specially tailored model. On the other hand its prime application domain being the educational context lead to some difficulties in the clinical setting for which the knowledge-based system was being developed. For example, the normal setting of Macpee, corresponding to a young 70 kg healthy male, requires adjustment if used in a clinical setting. Furthermore, Macpee cannot accept laboratory and clinical data directly as input and hence cannot provide a simulation of a specific patient. The approach adopted was to aggregate the infinite spectrum of possible patient conditions into 15 diagnostic states. The patient data are used to diagnose the patient and the diagnostic state then defines the values to which the parameters of Macpee need to be adjusted. It is thus assumed that the model behaviour is representative of the diagnostic class to which the patient has been assigned; and hence the recommendation made by the treatment module can be tried out on the model to justify this therapeutic action.

The communication between the knowledge-base and the model is achieved by means of a text file. Hence, as the program goes through its line of reasoning, the diagnosis and management strategies are translated into instructions that Macpee can understand, and stored in the appropriate file.

Macpee had to be altered in so that it would read instructions from the file rather than the keyboard. Currently, when the control of the program is passed to Macpee, it will look for the instruction file first and execute all the commands. If the last instruction in the file is not to return the control back to the knowledge-base, the control is passed to the keyboard so that the user may experiment with Macpee. Once Macpee is aborted, the control is returned to the knowledge-base.

Example of some of the rules to change the parameters of Macpee are given below:

IF cardiovascular performance is diminished
and oedema is present
THEN reduce the cardiovascular contractility

IF patient has history of heart failure
and reduce cardiovasular contractility
THEN change the cardiovascular contractility
to 50%

IF change the cardiovascular contractility 50%
THEN OPEN the interface file
WRITE instruction
CLOSE interface file.

In order to keep the rule base small, **change-parameter** is defined as a predicate with two arguments: the name of the parameter and its new value. Therefore, the last rule above is a rule with its arguments being assigned to cardiovascular-contractility and the value 50%. A dictionary is formed of the name of changeable parameters and their corresponding factor number that Macpee recognises.

There was an implementation problem with fluid infusion in Macpee that was rectified. A combination of fluids could not be administered rather the last fluid listed would be simulated.

4.5.8 Implementation

The prototype was developed on an IBM PC with 640K RAM, running PC DOS. Macpee is written in FORTRAN 77, whilst the other modules, including the user interface, were developed in LPA-PROLOG version 1.4.

Due to the heavy requirements of Macpee and Prolog on RAM, and also the limitation of MS-DOS (maximum of 640K), both programs do not fit in the memory at the same time. This meant the RAM must be cleared before being able to load Macpee, which is time consuming and cumbersome.

4.6 DIAGNOSTIC MODULE REVISITED (MK II)

As mentioned previously, the diagnostic module cannot handle the presence of multiple disorders. The majority of the patients admitted to an intensive care unit either suffer from a number of disorders or develop secondary conditions. A knowledge-based system should, therefore, be able to cope with these patients if it is to be used in that environment.

The diagnostic module needed major reconstruction so that it could consider a patient in a more clinical manner. The physiological system was broken down to its sub-systems and rules developed to assess the condition of the patient within the context of each sub-system. This is very similar to the approach taken by the treatment module.

The output of the system is an assessment of various sub-systems, rather than a definite diagnosis. More clinical data is required by the system to cope with various categories of patients routinely admitted to an intensive care unit.

The program obtains a history of the patient by asking questions with regard to the history of heart failure, acute heart disorders and drugs that are currently being taken. The parameters of Macpee are instructed to be altered to correspond to these effects. In this way it is cosidered that the model simulates the condition of the patient more closely.

An example of the new rules is given below:

IF patient is NOT on diuretics
and Diastolic pressure < 90
and Urine output < 41
and Urine-Na < 10
and Temp-difference (core & periphery) High
THEN Fluid-overload Functional.

The above rule has established that fluid overload is due to a functional failure rather than an excessive cosumption or over administration of fluids. This is a typical assessment of the state of the patient which is reported to the user and used to change the parameters of the model. The rule to change parameters of Macpee is identical to the example given in section 4.5.4.

The focus of attention in this version of the diagnostic module is on the key factors (variables) that are associated with each disorder. Each disorder has a

particular effect on each subsystem which is characteristic of it. The problem-solution then was to identify these effects in isolation. The intention was to identify disorders rather than classify them.

The treatment module was subsequently modified and expanded to generate treatment recommendations based on the assessment of the diagnostic module. Therefore, the original configuration of the prototype (figure 4.1.) was realised, and the link between the diagnostic and treatment modules was established.

4.7 EVALUATION

One of the requirements of the DHSS was that the prototype should undergo formal evaluation. The process of evaluation should formally highlight the weak points of the prototype as well as its strong points. A formal evaluation would also allow the methodology and approach adopted by the group to be assessed.

4.7.1 Introduction

Systems are continually being evaluated, whether consciously or not. Designing and implementing knowledge-based systems involves constant evaluation of the progress by considering questions such as:

* Is the knowledge representation scheme adequate or does it need to be extended or modified?

* Is the system coming up with right answers and for the right reasons?

* Is the embedded knowledge consistent with the expert?

* Is it easy for users to interact with the system?

* What facilities and capabilities do users need?

Feedback from users, expert collaborators, and the system builders suggests improvements that may be incorporated into later versions. Evaluations pervade the system-building process and are crucial for improving system design and performance. Each time a rule in the knowledge-base is changed, added, or deleted, everytime the code of reasoning program is modified or extended or the knowledge representation scheme is refined, action has been taken in response to an informal evaluation.

One reason for the present difficulty in evaluating knowledge-based systems is that human experts are seldom evaluated objectively. Evaluations are very easily misinterpreted, and it is therefore extremely important to carry out evaluating at appropriate stages in the

system's development, to clarify exactly what is being evaluated, and to interpret the results correctly.

Domain experts involved in the construction of the knowledge-based systems are concerned primarily with the embedded domain knowledge and how it is used by the program. Thus the experts repeatedly perform both static and dynamic evaluations. In static evaluation they compare the performance of the knowledge-based system with their own conceptual model, looking for consistency and completeness. In dynamic evaluation they compare the system's line of reasoning and its conclusions in a specific case, with their own. Thus, the knowledge acquisition process is intimately linked with ongoing evaluation by the domain experts. Furthermore, the ongoing evaluation helps them structure and understand better both their domain and their own expertise.

The ultimate criterion of success is whether a knowledge-based system is actually used for expert consultation by individuals other than the system's developers. Only a few expert systems have reached this stage. A key ingredient of success is involving eventual users in evaluation of the system as it is being built. Without a clear understanding of the ultimate user's needs and requirements, system builders may fail to provide crucial capabilities and, consequently, the system may have limited utility. Involving end users in

the development process can generate user interest in the system as a potential tool for their own use.

There is a clear need for formalism in the process of evaluation. In designing an evaluation, one must be aware of its purpose: who is it for, exactly what is being evaluated and what one hopes to gain from the experiment.

4.7.2 Problems in the Evaluation of Medical Diagnosis Systems

Medical diagnosis, by its nature, suffers from uncertainties at different levels. These could be due to lack of a full understanding of the clinical domain as well as an inability on the part of clinicians to explain their processes of reasoning. Medical diagnosis systems cannot escape from these uncertainties and should therefore, be evaluated with these constraints in mind. Chanderasekaran (1983) provides some guidelines:

a) Success/Failure- When evaluating performance of complex systems, especially at a development stage, simple "success" vs. "failure" evaluations based on the final answer may be insufficient because they do not take into account the possibility of very acceptable intermediate performance. As pointed out by Yu et al. (1979), "A complex

reasoning program must be judged by the accuracy of its intermediate conclusion as well as its final decision".

b) "Correct" answers may be unknown- Often there are no "correct" answers, since expert clinicians may disagree among themselves.

c) Small sample size problem- Performance of the system in "rare" diseases cannot often be reliably evaluated due to the generally small sample size of the available cases.

d) Matching distribution of clinical practice- Without some knowledge about the distribution of types of cases that a system will need to confront, the results of evaluation cannot be suitably interpreted. For instance, suppose the system is very efficient in solving most of the "common" occurrences of diseases in an area of clinical medicine, and relatively poor in solving rare or "difficult" cases. If the difficult cases were to be chosen because they are "interesting" as test cases, the statistical evaluation of the system might not represent its performance in a real clinical setting. A solution to this situation is to require that cases be selected as representative of the target clinical setting.

4.7.3 Evaluation of the Prototype

The system under development has been informally evaluated from the start of the project. A semi-formal evaluation took place when the system was installed at West Middlesex University Hospital, London. This provided an opportunaty to assess user requirements and whether the limited knowledge represented in the system was adequate to deal with real life problems.

The initial feedback was the inadequacy of the database in handling and retrieving data from the patient files. There was no built-in facility to differentiate between data input at different times. To cope with this problem an indexing method was devised, which has been described previously.

A more fundamental and serious shortcoming of the system was the performance of the diagnostic module. Although it was capable of diagnosing well-defined disorders, it was unable to identify multiple disorders. This resulted in the review of the knowledge representation schema, which has also been described above.

A formal evaluation of the prototype is being planned. The focus of attention will be on the performance of the diagnostic module, treatment module

and the usability of the prototype (i.e. convenient interaction).

Three kinds of data will be put to the system, historical, hypothetical, and data available on the existing patients at the intensive care unit.

The knowledge-based system under development has two major distinct components; the knowledge-base itself and the mathematical model. It should be emphasised that although both components will be evaluated as a complete system, only the knowledge-base will be modified by the group.

4.8 SUMMARY

In this chapter, a detailed description of the two versions of the developed prototype was presented. Problems encountered with the first version (MK I) and hence, the motives for the development of the second version (MK II) were introduced. The necessity for evaluation as an ongoing, integral part of the development process was highlighted. In the next chapter, some clinical cases are put to the two versions of the prototype and their responses are compared and contrasted.

CHAPTER 5

[redacted]

(Alexis de Tocqueville)

5.0 SIMULATIONS OF THE KNOWLEDGE-BASE

The two versions of the prototype which have different methods of knowledge representation- list of expected patterns (MK I) and rule-based (MK II)- were described in detail in the previous chapter. This chapter demonstrates how the two versions respond to four clinical cases, three of which illustrate single disorders and the fourth multiple disorders.

The objective is to demonstrate the reasoning mechanism of the two versions as well as highlighting the capability of the second version to diagnose and treat cases with multiple disorders.

5.1 PROBLEM 1

Plasma Na	123	mmol/l
Plasma Urea	9	mmol/l
Creatinine	80	umol/l
Urine Na	50	mmol/l
Diastolic Pressure	70	mmHg
CVP	5	cm-H_2O
Urine Output	35	ml/h

Table 5.1 Reference Ranges Used by System for Classification of Data.

	LOW	NORMAL	NORMAL	HIGH	
Plasma Na	130	135	145	150	mmol/l
Plasma K	3.0	3.49	4.51	5.0	mmol/l
Plasma Urea	3	3	6	15	mmol/l
Haemoglobin	13	13	18	18	g/dl
Creatinine	60	60	120	160	umol/l
Albumin	20	35	50	50	g/l
Urine Osmolality	250	250	400	400	mmol/kg
Urine Na	10	10	20	20	mmol/l
Urine Albumin	0	0	0.01	0.01	g/l
Diastolic Pressure	60	70	90	90	mmHg
Temp. Difference	0	0	1	1	
CVP	3	3	8	8	cmH_2O
PCWP	6	6	12	12	mmHg
Urine Output	40	40	200	200	ml/h
Urine K	60	60	80	80	mmol/24h

5.1.1 Response of MK I

The system classifies the above data into appropriate qualitative categories (the reference ranges used are shown in table 5.1.) and produces the following correct diagnosis:

Addison's Disease

The treatment module checks the cardiovascular data and prompts:

No Cardiovascular Abnormalities Detected

Administer The Following:

Colloid	**500 ml**
NaCl	**2000 ml**

COMMENTS:

The prototype recognises the pattern and matches it to one of the represented disease-states. The treatment module acts independently of the established diagnosis and provides an acceptable strategy.

5.1.2 Respose of MK II

The system starts by asking questions regarding the history of cardiac problems and any drugs being administered. It establishes that there is no history of Myocardial Infarction and adds this information to its data base. Cardiovascular and renal data are then considered and the prototype prompts:

The high urea/creatinine ratio together with high urine sodium excretion suggests renal damage or Addison's disease.

The treatment module prescribes the following:

Colloid 500 ml

NaCl 2000 ml

COMMENTS:

The system indicates some causes of impaired renal sodium conservation and suggests appropriate treatment. The diagnosis is based on a mechanistic explanation of the abnormal data.

5.2 PROBLEM 2

Plasma Na	137	mmol/l
Plasma Urea	3	mmol/l
Creatinine	75	umol/l
Urine Na	45	mmol/l
Diastolic Pressure	75	mmHg
CVP	9	cm-H_2O
PCWP	30	mmHg

5.2.1 Response of MK I

In this example, the prototype finds two states that do not contradict the above data and hence produces the following differential diagnosis:

Normal Plasma-Na with Overload

Overtransfusion of Blood

The treatment module responds in the following manner:

Fluid overload, could be with or without Congestive Cardiac Failure (CCF). Treatment is to induce fluid deficit and

consider inotropes and/or veno-vaso dilating drugs if CCF is present.

Administer:

Dextrose 5% 1000 ml

COMMENTS:

The two diagnoses are quite similar, that is, the patient has fluid overload. The treatment module confirms the diagnosis but cannot establish the cause. The suggested treatment is to restrict fluid intake and to administer appropriate drugs if heart failure is considered to be a real possibility by the user.

5.2.2 Response of MK II

The system asks questions regarding the history of any previous heart problems. A search through its rules results in the following assessment:

Fluid overload actual. Overtransfusion or overinfusion. Oedema may be present.

Treatment:

Fluid restriction and adminster:

Dextrose/Saline 1500 ml

COMMENTS:

Because the prototype asks questions regarding the history of the cardiovascular state, it is able to give a more definite diagnosis and provide clues to the cause of abnormality. It also suggests that the presence of oedema is a

strong possibility. The prescribed treatment is to restrict fluid intake and maintain the level of plasma sodium.

5.3 PROBLEM 3

Plasma Na	170	mmol/l
Plasma Urea	8	mmol/l
Creatinine	130	umol/l
Urine Na	5	mmol/l
Diastolic Pressure	75	mmHg
CVP	1	cm-H_2O

5.3.1 Response of MK I

Only one disease-state is found not to be contradictory to the data. The diagnosis is:

Hypernatraemia due to hypotonic fluid loss

The treatment module response is:

Administer colloid to raise CVP.

Is there evidence of negative fluid balance? ==> *Yes*

Administer:

Dextrose 3500 ml

COMMENTS:

Because of lack of interaction between the diagnostic and treatment modules, the system asks a question the answer to which should be clear from the established diagnosis.

5.3.2 Response of MK II

The diagnosis produced by the system is:

Fluid deficit actual.

Hypernatraemia and therefore water loss.

Treatment:

Raise CVP by giving Dextrose 5%

Stop when fluid causes CVP to rise above normal.

Dextrose 3500 ml.

COMMENTS:

The diagnosis is identical to the diagnosis of MK I, however, the system does not need to ask complementary questions because of the communication between the diagnostic and treatment modules. The treatment is slightly different but essentially the same approach is taken.

5.4 PROBLEM 4

Plasma Na	120	mmol/l
Plasma Urea	35	mmol/l
Creatinine	500	umol/l
Urine Na	30	mmol/l
CVP	10	cm-H_2O
Urine Output	100	ml/h
Albumin	20	g/l
Urine Albumin	8	g/d

5.4.1 Response of MK I

The matched disease-state is:

Water overload with renal failure

Treatment is:

Fluid overload with or without CCF.
Restrict water intake to 20 ml/hr.

COMMENTS:

The prototype correctly diagnoses two of the abnormalities present but fails to recognise low levels of plasma Albumin. The recommended treatment is adequate for the treatment of water overload but it fails to recommend any strategies to treat the renal failure.

5.4.2 Response of MK II

The prototype checks the history of Myocardial Infarction and looks for nephrotic syndrome. The output is:

The patient is nephrotic.
Fluid overload actual.
Impaired renal function; moderate-severe renal failure.
This patient has hyponatraemia with fluid overload, renal failure and the nephrotic syndrome.

Treatment is:

Fluid restriction.
Renal failure regime.
20 ml/hr water.

If oedema present 20ml/hr colloid + urine output until sodium becomes normal.

Colloid/NaCl 1000 ml

Dextrose 5% 500 ml

COMMENTS:

This version manages to detect all the abnormalities present and provides a summary at the end of the assessment. The treatment is to take action to correct all the detected abnormalities.

5.5 CONCLUSIONS

A number of cases were presented to both versions of the prototype. It is clear that MK II is a more clinically orientated version with adequate communication between its modules. MK II considers the data in different contexts. This ensures that all posssibilities for interpretation of the data are considered.

The last example highlights the inability of the first version to interpret all the abnormal data and the recommended treatment is also limited. MK II breaks the problem into its sub-components and is therefore able to give a correct assessment that explains all the abnormal data as well as recommending appropriate measures to correct them.

The two versions of the prototype will be discussed in more detail in the next chapter.

CHAPTER 6

[redacted]

(Friedrich Nietzsche)

6.0 DISCUSSION

In the previous chapter, the prototype under development was described in some detail. In this chapter the methodology undertaken, its strengths and weaknesses will be fully discussed.

6.1 INTRODUCTION

As mentioned previously, clinicians use various kinds of knowledge to diagnose and manage a patient depending on the complexity of the problem and the level of its development. A clinician confronts cases ranging from straightforward disorders which can be diagnosed quickly and efficiently using heuristic knowledge, to the more complicated in which s/he has to go back to first principles in order to understand the process.

A knowledge-based system should ideally contain knowledge similar to that which a clinician utilises in the process of clinical decision making. The knowledge available to a clinician comprises the principles of physiology and anatomy, some understanding of the pathophysiology of the specialised domain and the experience gained while applying this knowledge.

6.2 KNOWLEDGE REPRESENTATION

The developed prototype incorporates two knowledge sources, the knowledge-base and the mathematical model, Macpee. These knowledge sources are quite independent of each other, furthermore, knowledge is represented in a different manner in each case. They are discussed separately below.

6.2.1 Knowledge Base

The knowledge-base is comprised of symbolic representation of diseases and relevant management strategies. The knowledge is essentially associational; describing various patterns of signs and symptoms (manifestations) in relation to the underlying disorder.

The knowledge-base comprises rules describing various abnormalities and treatment strategies. These rules represent the pathophysiology of fluid and electrolyte metabolism heuristically. There are no rules defining the mechanism of action, the structure and the function of the components. Because of this shortcoming, it can only handle cases that are represented by their manifestations. It is unable to reason in terms of physiology or pathophysiology, and hence, incapable of handling disease-states that are not represented explicitly in the knowledge-base. Fluid-electrolyte is a complex domain with a large number of interactions between the subcomponents. A number of disorders can be present at the same time resulting in the "non-classic"

manifestations difficult to diagnose. Kuipers (1984), gives an interesting example of a patient suffering from *nephrotic syndrome* who is on a self-imposed low sodium diet. An important diagnostic finding is oedema, the swelling of ankles and wrists. However, this patient had no sign of oedema because of low salt intake. This would probably baffle a system whose knowledge is based purely on the association of manifestations with a disorder. This highlights the importance of the need for a more substantial knowledge-base to be present so that the underlying mechanism can be understood.

In this work an attempt has been made to tackle the problem of multiple disorders being present by concentrating on the key factors that are essential for the recognition of the presence of a disorder. A formal evaluation of the system in a clinical environment will allow further comments to be made on its effectiveness.

Clinical decision making is a difficult process to explain. Clinicians acquire through experience the ability to compile a long causal chain into an association (Kuipers, 1984). However, for a system to be acceptable, it should be able to explain the line of reasoning followed in an appropriate manner. The kind of explanation that can be generated using the knowledge-base described above is very shallow and unacceptable. Clinical users will not accept the decision of a system that is only able to produce explanations based on rules

that is only able to produce explanations based on rules that are triggered, especially if the rules represent the association between signs and symptoms with the underlying disorders.

A more acceptable explanation would be based on first principles (eg. Patil, 1981; Long, 1986). Recently, there has been enormous interest in "user modelling" explanation, where systems produce explanation based on the requirements of the user. For a review of various methods of explanation see Nicolosi (1988).

6.2.2 Model

The model implicitly represents the functional behaviour of fluid-electrolyte and the circulation. This representation is based on physical laws as well as physiological laws. Difference equations are used to represent relations and links between components. These links can be thought of as implicit causal links, simulating a cause-effect cycle. Models of this kind incorporate within them the results of experiments carried out on dogs and other animals to estimate the unknown parameter values in the equations. This is a kind of knowledge that is very valuable for an acceptable simulation but very difficult to represent explicitly. The mechanism of action of the domain, is then fairly well represented as far as realistic simulations are concerned. The main disadvantage of this kind of representation is their lack of transparency.

The model is quite capable of simulating multiple disorders. Various parameters can be changed to simulate a number of disorders in conjunction with each other. The model, however, does not take laboratory or clinical data as input and hence, it is difficult to fit a patient's data to the model. A mechanism had to be deviced by means of which the patient's data could be fitted to the model so that the model would simulate the condition of the patient. The knowledge-base was developed with this difficulty in mind, hence it acts as the front-end to the model, assessing the patient's condition and changing the appropriate parameters of Macpee to this effect.

6.2.3 The Combined System

The prototype therefore, incorporates two knowledge sources. One is based on heuristic knowledge (knowledge-base), the other on first principles and empirical knowledge (Macpee). The combination should be sufficient to meet the requirements of a clinical decision making system.

However, as the knowledge represented in the model is neither transparent nor explicit, it can only be used in a limited manner that is through simulations. Hence, the model is used as an external module with interaction being reduced to passing parameter values and receiving simulations.

The model is used to justify the recommended therapy regimen of the treatment module. As such, it is quite capable of simulating various plans and displaying the results in graphical and numerical form. Prompts are also displayed about how the patient is feeling when life threatening states are reached.

The two knowledge sources are developed from similar types of conceptual understanding of the mechanisms of action of the fluid-electrolyte and circulatory system. However, the groups developed them separately and without interaction with each other. This is an additional problem on top of the difficulties of implementation. There are areas in the domain of fluid-electrolyte and circulatory metabolism that are not fully understood and hence there is no uniform view about these problem areas.

6.3 REASONING

6.3.1 MK I

The first working prototype had two different reasoning mechanisms. The diagnostic module reasoning was a process of elimination of possible contenders for differential diagnosis. The disease-states represented were all potential cause of the manifestations unless there was at least one manifestation that was contrary to this hypothesis. This process meant that even though there was no evidence to support the presence of the remaining disorders in the differential diagnosis list,

there were no evidence to deny it either. The reasoning mechanism was therefore that of elimination rather than establishment of a diagnosis as the most likely cause of disorders.

This reasoning mechanism has the benefit of considering all the disease-states as a potential cause of the disorder without any bias or weighting. Hence, no disease-state would be rejected because of its uncommon occurrence.

There were a few disadvantages. Because disease-states were represented explicitly by their associational pattern of expected symptoms, a large number of disease-states had to be represented for it to be effective. The number would run towards infinity if multiple conditions were to be present. For example hyponatraemia and cardiac failure were represented by their expected patterns and if a patient was suffering from both conditions, a disease-state would have to be defined with the expected pattern. Another problem was the fact that the system would include disease-states within its differential diagnoses that were totally unrealistic but there would be no evidence to reject such a hypothesis. This problem was rectified by introducing additional constraints at the level of knowledge representation.

The treatment module performed much more efficiently and its recommendations were quite acceptable. It uses

various rules to assess the state of various sub-systems of the fluid-electrolyte and circulatory domain in order to propose appropriate suggestions. The sub-systems include the cardiovascular system and renal system. Various fluid therapy plans are generated using the above information together with the fluid balance state.

The interface between the knowledge-base and Macpee underwent two development phases. First the coding of Macpee was altered so that instructions could be given by default through a file rather than the keyboard. The second phase was to construct rules that would translate the output of the knowledge-base into instructions readable by Macpee and store them in appropriate order on to a file.

The output of the diagnostic module is used to tune Macpee to the state of the patient. Hence, the overall performance of the prototype is dependent on the diagnosis being correct. In order to reduce this dependency on the diagnostic module, the treatment module reasons quite independently of the diagnosis. In this way the performance of the two modules can be checked against each other.

6.3.2 MK II

An informal evaluation of the prototype indicated that the diagnostic module was capable of handling patients with single disorders subject to modification of

the constraints of the knowledge representation method. However, patients that are admitted routinely for intensive care usually suffer from more complicated conditions arising from a number of disorders. It was therefore decided to review the diagnostic module and as a result it was reconstructed.

The new diagnostic module uses rules to assess various relevant sub-systems. The assessments take the form of reports to the user on the condition of the sub-systems as well as translating the reports into instructions to change appropriate parameters of Macpee. In this version therefore, instead of a definite diagnosis, assessments are produced which results in a greater number of parameters of Macpee to be changed. It is considered that this approach will enable the model to track patient progress more closely.

This approach closely resembles the approach of the treatment module. Efforts were made to integrate the two modules. The final prototype produces treatment recommendations based on the assessment of the diagnostic module.

One of the problems that computer systems in medicine are faced with is the evolution of patients over time. That is, the patient's condition will change (for better or worse) either because of intervention by clinicians or as the result of the physiological system

correcting itself and compensating for deranged effects, or both.

The physiological model, Macpee, is a rich source of knowledge about the physiology and pathophysiology of the fluid-electrolyte and circulatory systems. As such, it is capable of tracking a patient over time. Macpee can also be used to justify recommended therapies. Time is therefore represented in an abstract manner at the therapeutic level.

The output of Macpee is numerical and graphical in form. It must be noted that to justify a recommended therapy, a change in the right direction of appropriate parameters from the abnormal state is considered to be sufficient and necessary, rather than a number to number correspondence. That is, the patient is not expected to have the same quantitative parameter values, having gone through the recommended therapy, but rather the same shift in direction from the abnormal state.

6.4 LIMITATIONS

The prototype suffers from various limitations. Some are the inevitable reflection of uncertainties in the medical domain whilst others are the result of implementation and conceptual limitations.

Patient data are classified into various ranges before used for processing. The classification is

carried out by comparing the data with a set of reference ranges. This creates a problem when data fall close to the threshold of adjoining ranges. The reference ranges are based on the expected values for "typical" patients. In reality, a patient may have a large rise in a particular parameter but still fall into the "normal" range. This factor undermines to some extent the effectiveness of the diagnostic module.

The prototype's reasoning process is in effect an "open loop". That is, it makes diagnosis, recommends appropriate therapy regimes and simulates these on the model, but no assessment is made from the result of the simulation to modify the therapeutic measures if needed.

The two knowledge sources were developed independently, although the knowledge-base was developed so that it would compensate for some of the deficiencies of Macpee. However, the physiological and pathophysiological knowledge represented have been developed by different groups at different times. There may be some conflict as a result.

It is of utmost importance for a system in a high risk environment like the intensive care unit to provide explanation of its decision to the user for it to be acceptable. Although the model is used to justify recommended action, this justification cannot replace explanation. The protype has no explanation facilities.

Because of the implicit and heuristic method of knowledge representation, the system as it stands can only be made to generate explanation at a shallow level. For instance, a particular action is recommended because a particular pattern of manifestation is present.

6.5 RECOMMENDATIONS FOR FUTURE WORK

Further development can be carried out to overcome some of the deficiencies of the system mentioned above.

For instance, in tackling the problem of ranges, one piece of data should not be considered in isolation but the overall trend over an appropriate time scale must be considered.

The reasoning loop could be closed by a qualitative assessment of the key parameters after simulation, so that new actions can be recommended.

The other limitations are more fundamental and require major reconstruction.

A coherent knowledge-base should be developed based on explicit representation of the cause-effect cycle, where this is known, and empirical knowledge, to clarify the ambiguities of uncertain areas. Kunz (1984), Widman (1986) and Long (1986), are three examples of this approach. This coherent knowledge-base, could then be used to generate diagnosis, therapy, prediction and most

importantly explanation. Because of the explicit representation of the knowledge, the system would be able to produce an acceptable explanation to justify its reasoning mechanism.

CHAPTER 7

[redacted]

(Albert Einstein)

CONCLUSIONS

The objectives of the project were to study the feasibility of coupling a dynamic mathematical model to a knowledge-based system, and to assess the benefits of such coupling. A prototypical approach was taken to achieve the objectives. A prototype was developed and is to be used in the Intensive Care Unit at the Royal Free Hospital, London, to diagnose, treat and predict the state of patients suffering from fluid-electrolyte and circulatory disorders.

The work discussed in this thesis has made a number of contributions both to systems science and clinical medicine.

The contributions made from the systems science perspective are:

* Through the development of the prototype, it has been shown that coupling of a dynamic mathematical model to a knowledge-based system can be achieved.

* This case study has highlighted the benefits of such coupling, and equally revealed the limitations. Time was represented at the therapeutic level, enabling the system to predict future state of the patient. It was not possible, due to the limitations of the model, to represent time at the diagnostic level.

* The limitations provided clues as how better to proceed with coupling. Furthermore, the limitations provided guidelines as how to proceed with the process of a complete integration of a mathematical model with a knowledge-based system.

From a clinical perspective the contributions of this work are:

* A prototype was developed which encapsulated expert knowledge within it in an accessible manner, to assist nurses as well as clinicians in the management of patients in a Intensive Care Unit.

* A dynamic mathematical model was coupled to a knowledge-based system and thus a temporal element was introduced at the therapeutic level of the developed prototype. This allows the system to track the patient over time and

predict changes with regard to the therapeutic measures.

* The key parameters used to diagnose and manage patients were identified by studying the data processing mechanism of the clinicians.

* As a result of the knowledge acquisition and elicitation processes, the clinical decision making process became more structured.

* It became apparent that clinicians use various types of knowledge at various levels of detail at different times according to the complexity of the problem at hand. A knowledge-based system should have similar knowledge represented within it, taking quantitative and qualitative forms, representing causal and heuristic links. Such a knowledge-base is rich enough to be a source for reasoning and explanation.

REFERENCES

Bobrow D.G., and Winograd T., (1977).
An overview of KRL, a knowledge representation language.
Cognitive Science vol. 1, pp. 3-46.

Brachman R.J., (1977).
What is in a concept: structural foundation for semantic networks.
International Journal of Man-Machine Studies vol.9, pp. 127-152.

Bratko I., Mozetic I., and Lavrac N., (1986).
Automatic synthesis and compression of cardiological knowledge.
In Michie D., and Bratko I., (Eds.), *Expert Systems: Automating Knowledge Acquisition*. Addison-Wesley, Reading Massachussets.

Carson E.R., Cobelli C., and Finkelstein L., (1983).
Mathematical modelling of metabolic and endocrine systems: model formulation, identification and validation.
Wiley, New York.

Chandrasekaran B., (1983).
On evaluating AI medical diagnosis.
AI Magazine, Summer issue, 34-38.

Cramp D.G., (1975).
The measurement and modelling of blood glucose dynamics in man.
PhD. thesis, London: Royal Free Hospital School of Medicine, University of London.

Cramp D.G., Carson E.R., and Leaning M.S., (1985).
Some design features for a user-friendly computer-aided decision support system incorporating mathematical models.
In: de Lotto I., and Stefanelli M., (Eds.): *Artificial Intelligence in Medicine*, Elsevier, Amsterdam, pp. 11-20.

Dickinson C.J., Ingram D., and Ahmed K., (1987).
MACPEE graphical version: a simulation of heart, peripheral circulation, kidneys, body fluids, electrolytes and hormones.
IRL Press, Oxford.

Feigenbaum E.A., (1977).
The art of artificial intelligence.
Proc. IJCAI-77, pp. 1014-1029.

Finkelstein L., and Carson E.R., (1985).
Mathematical modelling of dynamic biological systems.
2nd. ed. Research Studies Press, Letchworth.

Flood R.L., (1985).
Quantitative modelling of the fluid-electrolyte, acid-base balance for clinical application.
PhD. thesis, City University, London.

Flood R.L., Carson E.R., and Cramp D.G., (1986).
The control of body fluid volume: mathematical models of the regulatory process.
In Cramp D.G., and Carson E.R. (Eds): *The Circulatory System*, Croom Helm, London, pp. 143-221.

Genesereth M., Greiner R., Smith D., (1980).
MRS manual.
Technical Report, Stanford University, December 1980. Stanford Heuristic Programming Project Memo HPP-80-24.

Guyton A.C., Coleman T.G., and Granger H.J., (1972).
Circulation: overall regulation.
Annu. Rev. Phyiol., 34, pp. 13-46.

Guyton A.C., Coleman T.G., Manning R.D. Jr., and Hall J.E., (1984).
Some problems and solutions for modeling overall cardiovascular regulation.
Math. Biosci., **72**, 141.

Ikeda N., Marumo F., Shirtataka M., and Sato T., (1979).
A model of overall regulation of body fluids.
Ann. Biomed. Eng., **7**: 135-166.

Kowalik J.S., (1986).
Coupling symbolic and numerical computing in expert systems.
North Holland, Amsterdam.

Kowalik J.S., Chalfan K.M., Marcus R.I., and Skillman T.L., (1986).
Composite software systems.
Technical Report, Boeing Computer Services, Seattle, Washington.

Kuipers B., (1984).
Causal reasoning in medicine: analysis of a protocol.
Cognitive Science **8**, 363-385.

Kuipers B., (1985).
Qualitative simulation of mechanisms.
Cambridge, Massachusetts Institute of Technology, 1985. Technical report MIT/LCS/TM-274.

Kuipers B., (1986).
Qualitative simulation.
Artificial Intelligence **29**, pp. 289-338.

Kunz J.C., (1984).
Use of artificial intelligence and simple mathematics to analyze a physiological model.
Department of Computer Science, Stanford University, STAN-CS-84-1021.

Leaning M.S., (1980).
The validity and validation of mathematical models.
PhD. thesis, City University, London.

Leaning M.S., (1986).
A database-oriented architecture for large scale models: London, Medical School Computer Unit, Royal Free Hospital School of Medicine, 1986. Technical Report 86/3.

Leaning M.S., and Nicolosi E., (1986).
MODEL- software for knowledge-based modelling of compartmental systems.
Biomed. Meas. Infor. Contr. vol. 1, No. 4, pp. 171-181.

Long W.J., Naimi S., Criscitiello M.G., and Kurzrok S., (1986).
Reasoning about therapy from a physiological model.
MEDINFO 86, North Holland, Amsterdam, pp. 756-760.

Long W.J., Naimi S., Criscitiello M.G., Pauker S.G., and Szolovits, (1984).
An aid to physiological reasoning in the management of cardiovascular disease.
Computers in Cardiology Conference (1984), 3-6.

Long W.J., Naimi S., Criscitiello M.G., (1982).
A knowledge representation for reasoning about the management of heart failure.
Computers in Cardiology Conference (1982), pp. 373-376.

McAllester D.A., (1980).
An outlook on truth maintenance.
Technical Memorandum No. 551, Artificial Intelligence Laboratory, MIT, Mass.

Mason S.J., (1956).
Feedback theory- further properties of signal graphs.
Proceedings of the IRE, (1956), **44**, pp. 920-926.

Minsky M., (1975).
A framework for representing knowledge.
In P. Winston ed., *The psychology of computer vision*, McGraw-Hill, New York, pp. 211-280.

Mozetic I., Bratko I., and Lavrac N., (1984).
The derivation of medical knowledge from a qualitative model of the heart.
ISSEK Workshop 84, Bled, Yugoslavia, August 1984.

Newell A., (1969).
Heuristic programming: ill-structured problems.
In S. Aronofsky, ed., *Progress in operations research*, vol. 3, John Wiley and Sons, New York, pp. 360-414.

Newell A., and Simon H.A., (1963).
GPS: A program that simulates human thought.
In E.A. Feigenbaum and J.A. Feldman, eds., *Computers and Thought*, McGraw-Hill, New York, pp. 279-293.

Nicolosi E., (1988).
A critical review of approaches to computer-based explanation in knowledge-based systems.
Technical Report 88/1, April 1988, Royal Free Hospital School of Medicine, London.

Nicolosi E., (1986).
Critical review of extant systems combining artificial intelligence and classical mathematical techniques.
City University, London, Research Memorandum MIM/EN/3.

Patil R.S., (1981).
Causal representation of patient illness for electrolyte and acid-base diagnosis.
Rept. MIT/LCS/TR-267, Laboratory for Computer Science, Massachussetts Institute of Technology, Mass.

Pople H.E., (1977).
The formation of composite hypotheses in diagnostic problem solving.
Proc. 5th. Joint Int. Conf. Artif. Intel., Boston, MA., pp. 1030-1037.

Shamsolmaali A., Carson E.R., Collinson P.O., and Cramp D.G., (1988).
A knowledge-based system coupled to a mathematical model for interpretation of laboratory data.
IEEE Engineering in Medicine and Biology Magazine, (in press).

Shamsolmaali A., Carson E.R., Collinson P.O., and Cramp D.G., (1987).
A knowledge-based system for interpretation of laboratory data incorporating a mathematical model.
In *Proc. of Ninth Annual Conference of IEEE EMBS*, Boston Massachussets, November 1987.

Shortliffe E.H., (1976).
Computer based medical consultations: MYCIN.
Elsevier, New York.

Szolovits P., Hawkinson L.B., and Martin W.A., (1977).
An overview of OWL, a language for knowledge representation.

Rept. LCS-TM-86 Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, Mass.

Weiss S.M., Kulikowski S.M., Amarel S., and Safir A., (1978).
A model-based method for computer-aided medical decision making.
Artificial Intelligence **11**, 145-172.

Widman L.E., (1986).
Representation method for dynamic causal knowledge using semi-quantitative simulation.
MEDINFO 86, North Holland, Amsterdam, pp. 180-184.

Yu V.L., Buchanan B.G., Shortliffe E.H., Wrath S.M., Davis R., (1979).
Evaluating the performance of a computer-based consultant.
*Comput. Prog.*Biomed., **9**: 95-102.

Appendix 1 - examples of the menu facilities and window outputs of the prototype (pp.145-152)
has been removed for copyright reasons

Appendix 2 - listing of the developed prototype in LPA-PROLOG (pp.154-182)
has been removed for copyright reasons

APPENDIX 1

In this appendix some examples of the menu facilities and window outputs of the prototype are presented. The windows shown in pages 145-152, show the logical sequence of a typical interaction.

APPENDIX 2

In this appendix, the listing of the developed prototype in LPA-PROLOG is presented. This listing is subject to Crown Copyright.