MIT/LCS/TR-267

CAUSAL REPRESENTATION OF PATIENT ILLNESS FOR ELECTROLYTE AND ACID-BASE DIAGNOSIS

Ramesh S. Patil

This blank page was inserted to preserve pagination.

Causal Representation of Patient Illness for Electrolyte and Acid-Base Diagnosis

by

Ramesh S. Patil

October 1981

© Massachusetts Institute of Technology

This research was supported (in part) by the National Institutes of Health Grant No. 1 P01 LM 03374-0x from the National Library of Medicine.

Laboratory for Computer Science

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Cambridge Massachusetts 02139

Causal Representation of Patient Illness for Electrolyte and Acid-Base Diagnosis

by

Ramesh S. Patil

Submitted to the Department of Electrical Engineering and Computer Science on October 24, 1981 in partial fulfillment of the requirements for the Degree of Doctor of Philosophy

Abstract

Much of the medical knowledge in the first generation *AI in Medicine* programs is phenomenological; that is, it describes the associations among phenomena without knowledge of the underlying causal mechanisms. Although these AIM programs provide a good first approximation to the way clinicians reason, they fail to reproduce clinicians' reasoning based on a deeper understanding of the phenomena. More specifically, they do not deal with the knowledge of disease at different levels of detail, nor do they utilize causal relations to organize and explain the clinical facts and disease hypotheses. They also cannot deal with illnesses resulting from multiple diseases, especially when one disease alters the presentation of the others. Finally, they are unable to capture the notions of adequacy and parsimony that play such a large role in diagnosis. To explore these issues and rectify these deficiencies, we have undertaken the task of providing expert consultation for electrolyte and acid-base disturbances.

This thesis reports the implementation of ABEL, the diagnostic component of the consultation program. In it, we explore the problems of modeling the causal understanding of a patient's illness. We develop techniques for dealing with illness resulting from multiple interacting diseases. We describe a multi-level representation of causal knowledge, and explore issues of the aggregation of available case specific knowledge into concise summaries of the patient's illness. We discuss structural criteria for evaluating parsimony, coherence and adequacy of diagnostic explanations. We also explore some of the issues involved in information gathering and propose expectation-driven diagnostic planning as a means of improving it. Finally, we discuss the issues of explanation and justification of the program's understanding and argue that these facilities are crucial for acceptability of a consultation program.

Thesis supervisor: Peter Szolovits
Associate Professor of Electrical Engineering and Computer Science

Keywords: Causal Representation, Medical Diagnosis

CONTENTS

1. Introduction 8

1.1 Scope of Project 12
1.2 Choice of Domain 14
1.3 Brief review of Electrolyte and Acid-Base Disorders 15
1.4 Desiderata 19
1.4.1 Making a Correct Diagnosis 19
1.4.2 Continued Management of the Patient 19
1.4.3 Diagnostic Style 19
1.4.4 Mode of Interaction 20
1.4.5 Handling Discrepant Information 20
1.4.6 Explanation 21
1.5 Survey of AIM programs 22
1.5.1 Internist-I and Present Illness Program 23
1.5.2 CASNET/Glaucoma 25
1.5.3 Mycin 25
1.6 Outline of the Thesis 27

2. Examples 30

2.1 Example 1: Salmonellosis 30
2.2 Example 2: Vomiting and Salmonellosis 40

3. Representation of Medical Knowledge 43

3.1 Anatomical Knowledge 44
3.1.1 Anatomical Taxonomy 44
3.1.2 Material Flow Pathways 44
3.1.3 Anatomical Spaces 46
3.1.4 Miscellaneous Gross Anatomical Relations 48
3.2 Etiological Knowledge 49
3.3 Physiological Knowledge 50
3.4 Disease Knowledge 51
3.5 Causal Link 53
3.6 Multi-Level Causal Description 54

4. Structure Building Operations 61

4.1 Structure of a PSM 61
4.2 Initial Formulation 62
4.3 Some Definitions 63
4.4 Aggregation 65
4.4.1 Focal Aggregation 65
4.4.2 Causal Aggregation 66

4.5 Elaboration 69
4.5.1 Focal Elaboration 69
4.5.2 Causal Elaboration 70
4.6 Projection 72
4.7 Component Summation and Decomposition 72

5. Diagnostic Problem Formulation and Information Gathering 80

5.1 Global Diagnostic Cycle 82
5.2 Diagnostic Closure of a Hypothesis 83
5.3 Scoring the PSM 85
5.4 Scoring a Disease Hypothesis 85
5.5 Information Gathering Strategy 88

6. Examples Revisited 95

7. Conclusion 118

7.1 Summary 118
7.2 Limitations of ABEL and Future Directions 119

8. References 122

Appendix I. System Building Tool: XLMS 130

I.1 XLMS Concepts 130
I.2 The XLMS Interpreter 133

Appendix II. Explanation 134

II.1 Phrase Generator 134
II.2 Higher level explanations 135
II.3 Organizing causal Explanation 136

FIGURES

Fig. 1. A schematic for the overall system 13
Fig. 2. Carbonic acid - bicarbonate buffer equation 17
Fig. 3. Nomogram of acid-base disturbances 18
Fig. 4. Graphic depiction of the two Acid-Base hypotheses 31
Fig. 5. Comparision of hypotheses 1 & 2 at clinical level 33
Fig. 6. Comparision of hypotheses 1 & 2 at intermediate level 34
Fig. 7. The part-of hierarchy 45
Fig. 8. Material flow relations 46
Fig. 9. The containment relation 47
Fig. 10. Gross anatomical relations 48
Fig. 11. Etiological hierarchy 49
Fig. 12. Schematic description of a causal link 53
Fig. 13. Schematic description of the node structure 55
Fig. 14. Comparision of lower Gi fluid and of plasma 56
Fig. 15. The loss of electrolytes in lower Gi fluid 56
Fig. 16. Consequences of lower Gi loss described at next higher level 57
Fig. 17. Lower Gi loss expressed at an intermediate level 57
Fig. 18. Salmonellosis and its consequences expressed at the clinical level 58
Fig. 19. Layered description of link: salmonellosis causes dehydration 58
Fig. 20. Compiled link 59
Fig. 21. Node types 64
Fig. 22. Causal aggregation: fully unaccounted node 66
Fig. 23. Causal aggregation: fully accounted node 67
Fig. 24. Causal aggregation: partially accounted node 68
Fig. 25. An example of the elaboration process 71
Fig. 26. An example of component summation/decomposition 73
Fig. 27. Feedback loop represented using component summation 74
Fig. 28. Component summation/decomposition: Case 3 77
Fig. 28. (continued) 78
Fig. 29. An example of diagnostic closure 84
Fig. 30. An example of explained, unexplained and unaccounted findings 86
Fig. 31. Initial diagnostic closure for salmonellosis and acute renal failure 89
Fig. 32. Diagnostic closure separated for each possibility 90
Fig. 33. Diagnostic closures for each possibility projected forward 91
Fig. 34. The goal tree 92
Fig. 35. Serum electrolytes and the bar diagram 95
Fig. 36. Graphical description of acid-base disturbances 96
Fig. 37. Hypothesis 1 98
Fig. 38. Hypothesis 2 99
Fig. 39. aggregation of low-serum-K-1 100

Fig. 40. Aggregation of low-pH-1 100
Fig. 41. PSM for hypothesis 1 101
Fig. 42. PSM for hypothesis 2 102
Fig. 43. Diagnostic closure 1 103
Fig. 44. Diagnostic closure 2 104
Fig. 45. One complete cycle of diagnostic inquiry 105
Fig. 46. Diagnostic closure 3 106
Fig. 47. Diagnostic closure 4 107
Fig. 48. Diagnostic closure 5 107
Fig. 49. After all findings have been exhausted 108
Fig. 50. Hypothesis 1 with salmonellosis 109
Fig. 51. Hypothesis 2 with salmonellosis 110
Fig. 52. English description of the two hypotheses 112
Fig. 53. Initial PSM 114
Fig. 54. Revised PSM after vomiting is entered 115
Fig. 55. Final PSM after salmonellosis is introduced 116
Fig. 56. English text of the final explanation 117
Fig. 57. The XLMS hierarchy 131
Fig. 58. Feedback relation between acidemia and hypocapnia 137

Acknowledgments

I would like to express my thanks to all of the people who made this thesis possible:

Peter Szolovits, my thesis supervisor, for providing constant attention, guidance and for helping me formalize ideas when I could not see through the confusion;

William B. Schwartz for suggesting this thesis topic, teaching me about acid-base and electrolyte disturbances, and for articulating and refining many of the ideas presented in this thesis;

Randall Davis, for his helpful comments and suggestions on drafts of this document;

William Martin for introducing me to the area of knowledge based application systems and providing the intellectual environment in which these ideas germinated;

Lowell Hawkinson for developing XLMS and providing help in using it;

Bill Long for his encouragement and for being a very insightful sounding-board for ideas;

Glenn Burke for proofreading this document and for providing much needed system support;

Ken Church for many spirited arguments and discussions on this and other topics;

Bill Swartout for timely development of explanation methodology and for sharing his experience in using XLMS;

Stephen Pauker, Brian Smith, Ben Kuipers, Byron Davies, Howard Sherman, Harold Goldberger, Brij Masand, Gretchen Brown and other past and present members of Clinical Decision-Making group and Knowledge-Base System groups for providing the fertile environment and comradery needed to carry me through the thesis;

and finally, my wife Aruna and my family for bearing with me and providing constant encouragement without which this would not have been possible.

1. Introduction

In a 1970 article reviewing the role of emerging computer technology in medicine, Dr. William B. Schwartz notes

> *"If conventional remedies will not meet the demands imposed by society's broad commitment to extensions of health care, it is clear that new, even heretical, strategies must be devised. One such strategy will almost certainly involve exploitation of the computer as an "intellectual," "deductive" instrument — a consultant that is built into the very structure of the medical-care system and that augments or replaces many traditional activities of the physician. Already, several interesting steps have been taken in an attempt to extend the computer's role into this realm ... Indeed, it seems probable that in the not too distant future the physician and the computer will engage in frequent dialogue, the computer continuously taking note of history, physical findings, laboratory data, and the like, alerting the physician to the most probable diagnoses and suggesting the appropriate, safest course of action. One may hope that the computer, well equipped to store large volumes of information and ingeniously programmed to assist in decision making, will help free the physician to concentrate on the tasks that are uniquely human such as the application of bedside skills, the management of the emotional aspects of diseases, and the exercise of good judgment in the nonquantifiable areas of clinical care."*
>
> —Medicine and the Computer [Schwartz70, page 3]

The decade following these predictions saw a rapid growth in the field of *Artificial Intelligence in Medicine* (AIM) culminating in many promising programs, among which are Internist-I [Pople77], the Present Illness Program (PIP) [Pauker76], CASNET/Glaucoma [Weiss74], MYCIN [Shortliffe76] and Digitalis Therapy Advisor [Pauker76]. These programs represent the first efforts in the use of AI techniques in medical decision making, and can be characterized as the "first generation AIM programs". They have clearly demonstrated the feasibility and usefulness of AI techniques. Most of these programs have, in some trial, been judged to match expert physicians in their competence — this is indeed an outstanding achievement.

It is natural to question then: "What are the limits of their expertise? Why aren't we implementing these programs in many more areas of medicine and distributing them for clinical use?" To answer these questions we must take a deeper look at the programs and their performance. For example, although they are (on average) outstanding on their core set of anticipated applications, their performance can also be non-uniform; it tends to degrade rather ungracefully just outside their domain of expertise. Furthermore, these programs may be misled on difficult cases involving complex interactions or multiple disorders, even if these cases fall well within their domain of expertise. This leads to the inevitable conclusion that although the models of representation and deduction used in these programs are capable of providing moderate coverage over the area of application, they are nonetheless inadequate.

These observations have led to a re-evaluation of the techniques used in the first generation of AIM programs. The following insights have been gained by this evaluation. Firstly, the notion of causality is inadequately exploited in the first generation AIM programs [Smith79, Patil79, Pople81]. They do not utilize the structure provided by causal relations to organize the patient facts and disease hypotheses. They fail to capture the human notion that explanation should rest on a chain of cause-effect deduction. Secondly, they cannot deal with the effects of more than one disease present in a patient simultaneously, especially when one of the diseases alters the presentation of the others. Thirdly, they do not deal with the knowledge of a disease phenomenon at different levels of detail that a physician clearly has. Finally, the numeric belief measures as used by the first generation AIM programs do not provide adequate criteria for diagnostic reasoning. They are unable to capture notions such as adequacy and parsimony of a diagnostic possibility.

Much of the medical knowledge contained in the first generation AIM programs can be characterized as being phenomenological; that is, it describes the associations among phenomena without the mechanisms underlying the observed associations. Such phenomenological descriptions provide a good first approximation to the way physicians reason, but they fail to capture the physicians' reasoning in recognizing and dealing with the inherent discrepancies in their knowledge and with deduction based on deeper understanding of the phenomena. Contrasting the behavior of the first generation AIM programs and human experts, Szolovits notes:

> *"Consider what happens when two "rules of thumb" (as we may identify a bit of phenomenological knowledge in medicine) conflict. Every AIM program written so far evaluates that conflict by reducing it to a numerical judgment of likelihood (or certainty, belief, etc.) in the hypotheses it holds: Mycin computes a revised certainty factor, CASNET computes new weights, Internist computes new scores, and the digitalis program often computes a weighted sum of its observations to evaluate their joint effect. Thus, conflict, just as agreement, is reduced to a manipulation of strength of belief. Yet, by contrast, we believe that human experts make a much more powerful use of occasions where they detect conflict. They are not satisfied by a simple revision of their degree of belief in the hypotheses which they have previously held; they seek a deeper, more detailed understanding of the causes of the conflict they have detected. For it is just at such times of conflicting information that interesting new facets of the problem are visible. Conflicts provide the occasion for contemplating a needed re-interpretation of previously-accepted data, the addition of possible new disorders to the set of hypotheses under consideration, and the reformulation of hypotheses thus far loosely held into a more satisfying, cohesive whole. Much of human experts' ability to do these things depends on their knowledge of the domain in greater depth than what is typically needed to interpret simple cases not involving conflict."*
>
> —Artificial Intelligence and Medicine [Szolovits81a, pages 16-17]

To move beyond the sometimes fragile nature of today's programs, we believe that future AIM

programs must contain medical knowledge similar in depth of detail to that used by expert physicians. They must have anatomical, physiological and pathophysiological knowledge sufficiently inclusive in both breadth and detail to allow the expression of any knowledge or hypothesis that usefully arises in medical reasoning.

One of the important areas of medical diagnosis not adequately addressed by the first generation of AIM programs is the evaluation of the effect of more than one disease present in the patient simultaneously, especially when one of the diseases alters the presentation of the others. For example, let us consider a patient with diarrhea and vomiting leading to severe hypokalemia. Let us also suppose that we know about the diarrhea, but we are not aware of the vomiting. The observed hypokalemia is too severe to be properly accounted for by the diarrhea alone and therefore diarrhea cannot be considered as complete explanation for the observed hypokalemia. Given this fact, the diarrhea is either not responsible for hypokalemia or is only partly responsible. If the diarrhea is not responsible, then further reasoning is relatively easy: the problem simplifies to finding the actual cause. However, if diarrhea is partly responsible, a correct partitioning of the total observed hypokalemia between its two suspected causes is required, with a judgment of how well the two separate causes combined in the estimated proportions account for the patient's condition.[1] Notice how inadequate the simple assignment of a probability linking diarrhea and hypokalemia (as is commonly done in existing programs) is to capture the problem being described here.

The complexity and depth of medical knowledge is well recognized [Szolovits78]. Our understanding of medical expert reasoning suggests that an expert physician may have an understanding of a difficult case in terms of several levels of detail. At the shallowest level that understanding may be in terms of commonly occurring associations of syndromes and diseases, whereas at the deepest it may include a biochemical and pathophysiological interaction of abnormal findings. While it may be easier for a program to reason succinctly with medical knowledge artificially represented at a uniform level of detail,[2] a range of representations are needed to reason at a sophisticated level of competence [Patil81]. Unfortunately, very little attention has been paid to developing methods for coping with it. We take this as the central issue of this thesis.

---

1. All the previous programs allow the entire hypokalemia to be accounted for by diarrhea. In particular, Internist-I after allowing the hypokalemia to be accounted for by diarrhea will not allow hypokalemia to lend any support to the hypothesis of vomiting. PIP, on the other hand, will allow the entire hypokalemia to lend support to the hypothesis of vomiting as well as allowing it to be explained by diarrhea.
2. This does not pose serious difficulty in medical domains where the pathophysiology of diseases is not well developed, because in such a domain a physician relies primarily on his phenomenological knowledge. However, in a domain such as electrolyte and acid-base disturbances we are constantly faced with this problem because, on the one hand, the pathophysiology of the disturbances is well developed, and on the other, the pathophysiology of many of the diseases leading to these disturbances is relatively poorly understood.

Finally, we believe that the numerical (probabilistic or pseudo-probabilistic) belief measures as used by the first generation AIM programs for confirming diagnoses and guiding the diagnostic search do not provide adequate criteria for diagnostic reasoning. We believe that the evaluation methods for confirming a disease hypothesis should be different from the methods used for choosing the most promising disease hypothesis for diagnostic pursuit. A single criterion is almost certain to be inadequate for both these tasks. Furthermore, we believe that the probabilistic model by itself is inherently inadequate. For example, it fails to take into account the causal nature of the disease mechanisms, it fails to capture the notions of parsimony, coherence and adequacy of diagnostic explanation. In a study of problem solving activity of clinicians, Kassirer and Gorry note that

> *"In parallel with the processes by which the physicians built a case toward a final diagnosis, they assessed each diagnosis for coherence and adequacy. ...* (A diagnosis was considered *coherent* if all the symptoms and diseases contained in it were causally related to each other. A diagnosis was considered *adequate* when it accounted for all all known facts.) *... The physicians strove to attain parsimonious explanations for the findings and to accept a single explanation rather than make two or more diagnoses unless they were forced to do so."*
>
> — Clinical Problem Solving [Kassirer78, pages 249-250]

It is one of the central themes of this thesis that these problems cannot be avoided by relying solely on the numerical scoring mechanism; the programs must be provided with structural criteria to evaluate the disease hypotheses.

It is our belief that modeling the program's understanding of the patient's illness is crucial to capturing the expertise of clinicians. In this thesis, we will explore some of the issues involved in representing diagnosis. We will develop techniques for reconciling physiological reasoning with phenomenological reasoning and explore issues of aggregating all the available knowledge into concise summaries of the patient's illness. We will discuss structural criteria for evaluating parsimony, coherence and adequacy of diagnostic explanations. We will also explore some of the issues involved in information gathering and propose expectation-driven diagnostic planning as a means of improving it. Finally, we will discuss the issues relating to explanation and justification of the program's understanding.

To study these issues, we have chosen the task of providing expert consultation in cases of electrolyte and acid-base disturbances. The research presented in this thesis, the development of a program called ABEL (Acid-Base and ELectrolyte program), is a part of this overall effort. We describe a novel mechanism for representing ABEL's understanding of a patient's illness. This understanding is represented using a collection of data-structures called the *patient-specific models*(PSMs). Each PSM contains a hypothesis structure containing all known data about the patient, all currently held possible interpretations of these data, the causal interconnections among the known data and tenable hypotheses, and some indication of alternative interpretations

and their relevant evaluations. We describe the representation of medical knowledge and the processing strategies needed to enable ABEL to construct a PSM from the initial data presented to the program. The same representations and procedures are also used in revising the PSM during the process of diagnosis. Each PSM can be viewed as a partial explanation of the patient's illness.

Diagnostic problems are formulated by identifying the weaknesses and conflicts in the PSMs and by computing a *diagnostic closure* (DC) for each PSM. A DC associated with a PSM represents a collection of alternative completions of the partial explanation provided by the PSM. It brings together all the dependencies and expectations necessary for diagnostic inquiry, for evaluating real and apparent discrepancies in the incoming information, and for explaining the diagnostic alternatives under consideration. A plan for diagnostic inquiry is generated by decomposing a top level diagnostic problem into simple problems which can be directly solved by a question to the user. Finally, when an inquiry is completed, the new information gathered is assimilated into the PSMs and the diagnostic process is repeated.

1.1 Scope of Project

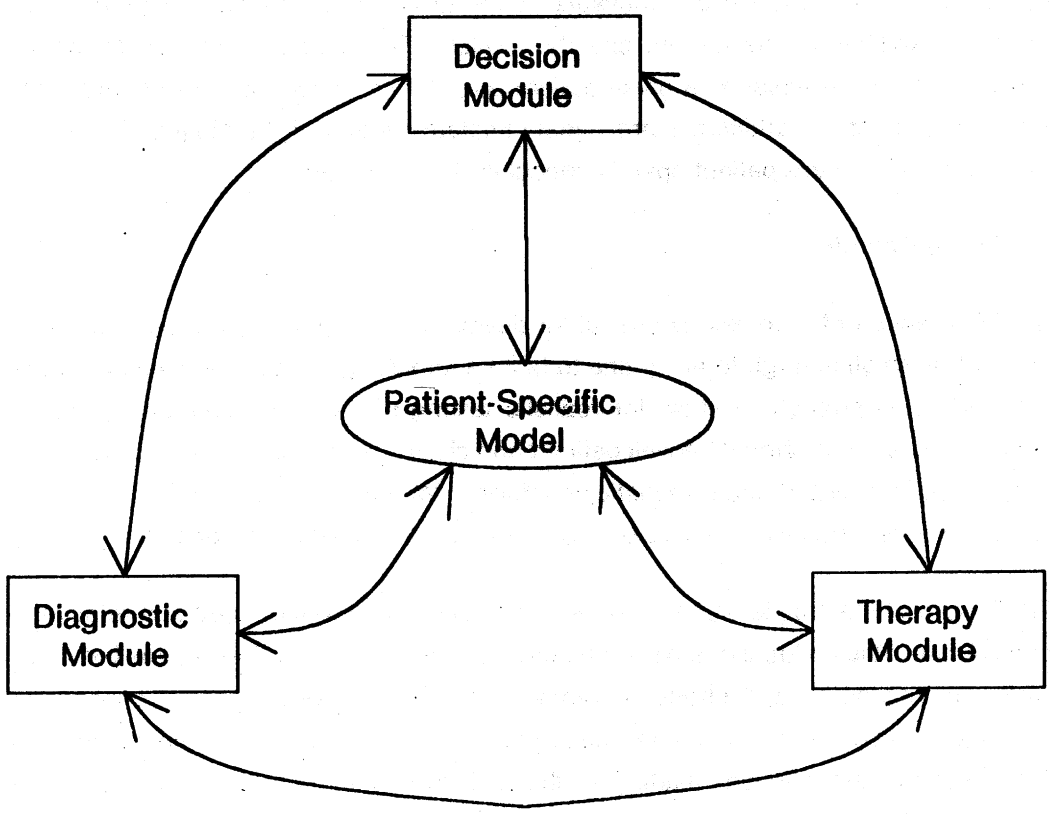
This thesis has three main objectives. The first is to develop a representation of causal medical knowledge. The second is to develop a case-specific "understanding" of illness. This understanding should be capable of describing subtle interactions between diseased and normal physiological mechanisms, and therapeutic interventions. The third is to develop a set of reasoning procedures to combine the aggregate phenomenological knowledge of disease associations with the detailed pathophysiological knowledge of disease processes. The first of these, the phenomenological knowledge, is necessary for efficient diagnostic exploration; the second, the pathophysiological knowledge, is necessary for proper understanding of a difficult case. The research reported in this thesis is conducted in the larger context of an Expert Consultant for Electrolyte and Acid-Base Disturbances [Patil79]. This section briefly reviews the organization of the overall system.

The objective of an expert medical consultant is to advise in the *proper management* of a patient. Proper management consists of collecting the relevant information about the patient, identifying the disease process(es) responsible for the patient's illness, and prescribing a proper course of action to correct the patient's condition. One of the complexities of this task is due to the fact that these subtasks do not have well defined boundaries. The patient may be presented to a clinician at different stages of a disease's evolution and treatment. During the course of management new information about the past history may become necessary as the diagnostic hypotheses evolve. The current diagnosis may depend on information that is presently unavailable. The disease itself may evolve through time, providing additional clues to its identity, or the response to certain therapeutic interventions may provide valuable diagnostic information. Finally, the patient's condition may require therapeutic intervention even before the diagnostic

issues can be reasonably resolved. Therefore, the next course of action must be chosen from a large range of alternatives. These alternatives may be broadly classified as *gathering information* (much of which may turn out to be irrelevant in the evolving clinical context), *ordering tests* (possibly involving expensive time delays and/or clinical costs), *waiting* for further development, *prescribing therapy* or some combination of the above. At every stage of consultation, the program must be able to choose between the alternative sets of actions with the patient's best interest in view. This can be achieved only by developing a program capable of forming a diagnosis, suggesting a therapy and making decisions. With this perspective we have embarked on the design of the Electrolyte and Acid-Base Consultant system. We have tried to separate and modularize different components of a physician's knowledge and expertise so as to be able to evaluate our understanding about each component and their interactions. This modularization should also allow us to further experiment with any component of the system without having to reimplement the entire program. A top level schematic for the overall system is shown in figure 1.

The Electrolyte and Acid-Base Consultant system consists of four major components: (1) the Global Decision Making component, (2) the Diagnosis component, (3) the Therapy component and (4) the Patient Specific Model. The patient specific model describes the physician's

Fig. 1. A schematic for the overall system



understanding of the state of the patient at any point during diagnosis and management; it is intended to be the central data structure which other components of the system may reason with. The global decision making component is the top level program which has the responsibility of calling the other programs with specific tasks. In general, the global decision program will call the diagnostic program with a task such as taking the initial history and elaborating some specific diagnosis. The diagnostic component then performs the specified task and reports the results to the main program. It also modifies the patient specific model to reflect the revised state of the patient. Similarly, if the global decision making program calls the therapy selection program, it attempts to formulate a set of alternate therapies for the patient along with a check list of items that must be tested before any specific therapy can be recommended. It also identifies information that will help discriminate between alternate therapy recommendations. Note that at every step the global decision maker can evaluate each of the possible sets of actions and choose the most desirable one. The decision making component will allow the program to make explicit the decision making that goes on in a physician's reasoning: is further diagnosis necessary, what treatment should be selected, should he wait before prescribing further treatment, can he choose some therapeutic action that would also provide diagnostic information making further diagnosis at this point unnecessary?

This thesis deals primarily with the development of the patient specific model which describes the program's understanding about the patient's illness. We have focused here because we believe that the level of expertise achievable by the program is inherently dependent upon the expressive capabilities of the patient specific model. The program can reason about subtle interactions between diseases in a given patient only if it can describe these interactions in the context of the patient. In addition a preliminary implementation of the diagnostic component to demonstrate the use of this patient-specific model is also discussed.

1.2 Choice of Domain

Careful selection of a domain is crucial for developing an application program. The domain chosen must be small enough to allow one to build a knowledge-base in a reasonable amount of time, and yet large enough to allow for realistic testing of the new ideas being implemented. Furthermore, the domain should be well defined and should lead to useful application, so that the program can be field-tested under realistic conditions. We have chosen the domain of electrolyte and acid-base disturbances as the test-bed for our theories of medical diagnosis.

The domain of electrolyte and acid-base disturbances is a well defined and relatively narrow area of medicine. It is an ideal domain for testing our theories about interactions between causal (physiological) reasoning and phenomenological (syndromic) reasoning, as on one hand the basic pathophysiology of the acid-base disturbances is well developed, and on the other, the pathophysiology of the diseases leading to these disturbances is relatively poorly understood. Thus constantly forcing us to develop reasoning mechanisms that can deal simultaneously with

well understood causal knowledge and poorly understood phenomenological knowledge. In addition, the feed-back nature of the electrolyte and acid-base homeostatic mechanism provides us, in a microcosm, with a variety of issues relating to "dynamic" systems that must be addressed in the management of a patient's illness.

Electrolyte and acid-base disturbances are a common complication of a large number of serious illnesses and medical interventions. In spite of their prevalence, this remains an area that most practicing physicians find somewhat difficult to deal with. This makes the field of acid-base disturbances an attractive domain for introducing expert computer consultant programs. One of the earliest programs for medical consultation [Bleich72] was in fact introduced in this very area.

Our primary concern, however is not with electrolyte and acid-base disturbances per se. Our basic purpose is to use this domain as a vehicle for evaluation of the existing techniques and development of new techniques for diagnosis and management of a patient's illness. In particular, in this thesis we will develop techniques for providing a coherent account of a patient's illness which incorporates the pathophysiological understanding of acid-base disturbances with the aggregate phenomenological understanding of the diseases causing these disturbances.

1.3 Brief review of Electrolyte and Acid-Base Disorders

In this section we briefly describe the electrolyte and acid-base disturbances. This section is not intended as a full review of the subject matter, but is presented here to provide the readers with a framework for understanding the medical examples used in this document. Each example used in the document is accompanied by an explanation of the relevant medical knowledge.

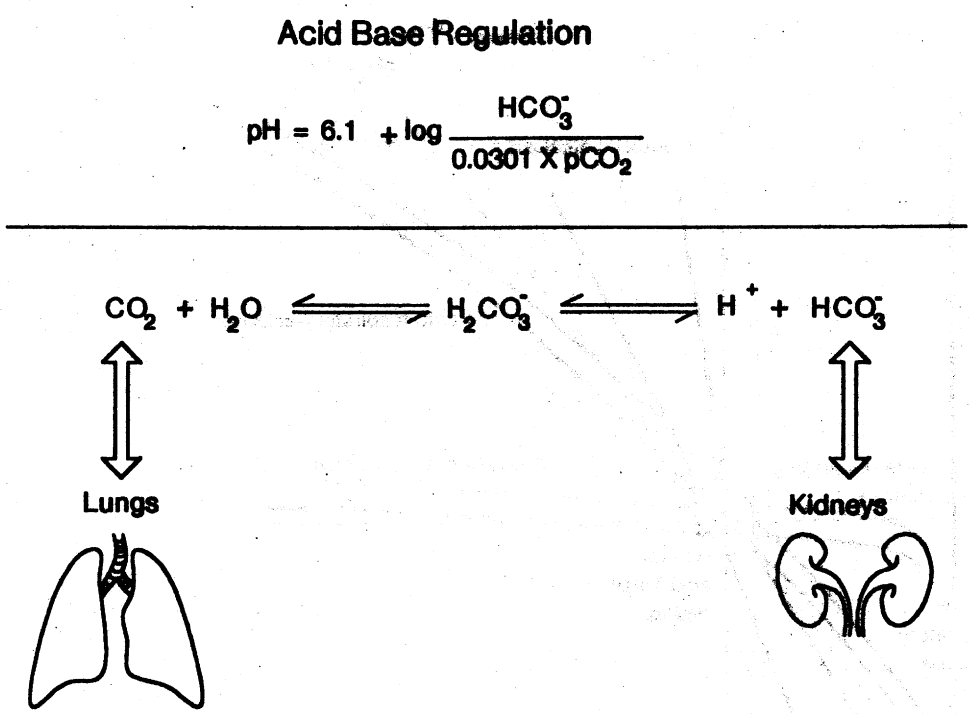
Fluid and electrolyte disturbances usually occur as complications of an underlying illness, therefore these disorders must be viewed not as isolated entities but in the context of the specific clinical settings in which they appear. As general background to the following discussion, it should be remembered that approximately 50 to 60 per cent of the body (by weight) consists of water distributed between the intracellular (within cells) and extracellular (outside cells) compartments. Water moves freely across cell boundaries, maintaining osmotic equilibrium between the different compartments. By contrast, owing to differences in their permeablity and active ionic pumps, the electrolytes are distributed in an asymmetric pattern, most of the ions in extracellular fluid consisting of sodium, chloride and bicarbonate and those in intracellular fluid of potassium and organic anions. Regulation of the external environment of cells, that is, the electrolyte concentration and acidity (pH) of the body fluids, is of primary importance. Perturbations in the regulation of this environment is the subject of electrolyte and acid-base disturbances.

The pH of the body fluids is regulated by three mechanisms: (1) the body buffers, (2) pulmonary regulation of the concentration of CO_2 in the body, and (3) renal excretion of acids and alkali. They act in a complementary fashion, first to minimize transient changes and then to correct any disturbances in acid-base balance by appropriate retention or excretion of hydrogen ions. To understand the mechanism of acid-base disturbances, it is instructive to consider the way in which the body deals with the normal daily acid load in maintaining a steady-state of acid-base equilibrium.

As food is oxidized to provide metabolic energy, both carbon dioxide (carbonic acid) and acids such as sulfuric and phosphoric acids are added to the extracellular fluid. They are immediately buffered to minimize the change in pH and transferred to the lungs and kidneys for excretion. Carbon dioxide is excreted almost entirely by the lungs while the other acids are excreted solely by the kidney. Bicarbonate is regenerated by the kidney as it excretes the excess acid, replenishing the bicarbonate stores that previously were depleted by the buffering of the dietary acid. From all these considerations it is evident that derangements in either the pulmonary or renal function, or the imposition of stresses that overwhelm normal regulatory mechanisms (such as vomiting, diarrhea, burns, etc.) can be expected to produce disturbances of acid-base equilibrium.

The equilibrium equation of the major buffer system in the extracellular fluid, the *carbonic acid — bicarbonate buffer system*, is shown in figure 2. This equation allows ready visualization of the directional changes that can be anticipated in both metabolic and respiratory disturbances of the acid-base equilibrium. For example, a primary reduction in bicarbonate concentration (*metabolic acidosis*) will cause the reaction to shift to the right, thus increasing hydrogen ion concentration, whereas a primary elevation in bicarbonate concentration (*metabolic alkalosis*) will cause the reaction to shift to the left, thus decreasing hydrogen ion concentration. Similarly, a primary rise in pCO_2 increases the hydrogen ion concentration (*respiratory acidosis*), and a fall has the reverse effect (*respiratory alkalosis*). However, the presentation of these disturbances is somewhat more complicated owing to the fact that the body reacts to these changes and attempts to compensate (in part) for the effect of these changes. Furthermore, different compensating mechanisms respond at different rates. A disturbance which has been properly compensated is called *compensated*, otherwise it is called *uncompensated*. The actual changes in the bicarbonate — carbonic acid concentrations in these disturbances is shown in figure 3. The nomogram of acid-base disturbances [Schwartz65, Cohen66] shown in figure 3 summarizes the normal physiologic response to the changes in HCO_3 and pCO_2 for each of the acid-base disturbances described above. For example, the nomogram shows that for a patient with adequately compensated metabolic acidosis and with serum concentration of HCO_3 of 15 meq/L the pCO_2 will be approximately 30 mmHg. The use of this nomogram for initial evaluation of a patient's acid-base state will be discussed later.

Fig. 2. Carbonic acid - bicarbonate buffer equation



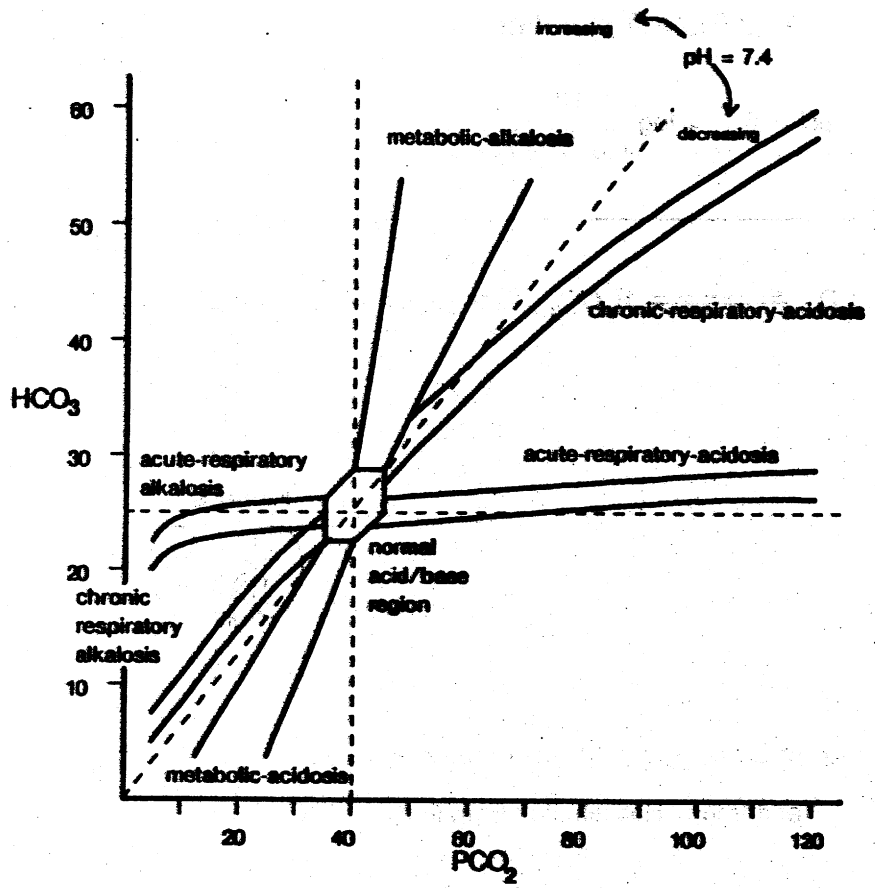
Regulation of Carbonic Acid / Bicarbonate buffer pair

Henderson - Hasselbalch Equation

The most frequently encountered clinical acid-base disorders occur as *single disorders* (also called simple disorders). The single disorders are: metabolic acidosis, metabolic alkalosis, respiratory acidosis, and respiratory alkalosis. There are, however, many clinical situations in which combinations of two or three disorders occur simultaneously, giving rise to *mixed disorders*. The recognition of mixed disorders is predicated upon a clear understanding of the pathophysiologic effects of simple disorders. To diagnose mixed disorders, one must know how each of the four simple disorders named above alter pH, pCO_2 and HCO_3 and the extent of renal or respiratory compensation that ought to occur for any given degree of primary disorder. However, since each of the disturbances can be caused by a variety of physiological states or diseases, the final differentiation between possible acid-base disorders must be made primarily on the basis of clinical information.

An important test in the diagnosis of electrolyte and acid-base disturbances is the laboratory analysis of a patient's blood sample. Also called the serum electrolytes, this test measures the concentrations of sodium (Na), potassium (K), chloride (Cl), and bicarbonate (HCO_3). Very often a test for concentration of creatinine is also made. This test does not, however, measure the

Fig. 3. Nomogram of acid-base disturbances



concentrations of anions such as phosphate, sulfate, proteins, and organic acids which are normally present in the blood in small amounts. The combined concentrations of these unmeasured anions is called the *anion gap*. The anion gap can be approximated by subtracting from the combined sodium and potassium concentrations the combined concentration of chloride and bicarbonate, an amount normally approximately 12 meq/L.

Determination of the anion gap is vital to the diagnosis and differentiation of metabolic acidosis. The anion gap differentiates metabolic acidosis into two categories: one with an increased anion gap and other with a normal anion gap. Metabolic acidosis with an increased anion gap is generally caused by increased production or impaired excretion of H^+ and unmeasured anions by the body. For example, diabetic ketoacidosis, in which the acidosis results from increased production of ketones. On the other hand normal anion gap acidosis is generally caused by loss of HCO_3. For example, diarrhea, in which HCO_3 rich gastrointestinal fluids are lost.

1.4 Desiderata

In this section we discuss some of the characteristics required of the program if it is to be useful and effective as an expert consultant. They also serve as guiding principles for designing and evaluating the program. They are included here to communicate our aspirations. The goals described below have not been fully realized by the research reported here, nor can they all be fully realized by the current state of AIM technology. These characteristics are:

1.4.1 Making a Correct Diagnosis

The primary responsibility of the diagnostic program is to make a *correct diagnosis*. Without fulfilling this criterion, the program offers little possibility of being clinically useful. Although the issues involved in the evaluation of the efficacy of diagnosis by a program (or by a clinician) are difficult and controversial, it is clear that the diagnosis arrived at by the program must be a reasonable and thorough diagnosis in the light of the available information. Furthermore, a distinction must be made between a *working* diagnosis and the *correct* diagnosis. In practice, a correct diagnosis is often impossible owing to the high cost (medical and economic) of the information necessary to achieve it. A criterion for deciding when a working diagnosis has been achieved (for the purpose of management of a patient) should weigh the costs of gathering further information in terms of morbidity, time and money vs. the benefits of better diagnosis in terms of an improved management plan and a more reliable prognosis. For example, in situations in which the management plan for each of the diagnostic possibilities is the same, attempts to distinguish between diagnostic alternatives does not have any immediate utility. Hence, the working diagnosis should be considered sufficient. It should, however, re-evaluate the diagnosis as new information becomes available from the evolution of the disease or from the patient's response to therapy.

1.4.2 Continued Management of the Patient

Typically, a patient is examined by a physician more than once. The interaction between the patient and the physician can be divided into the initial interaction and the follow-ups. The follow-up sessions are used by physicians in evaluating the management plans and in refining the working diagnosis. In the majority of cases, follow-up sessions are essential for the proper practice of medicine. Furthermore, the ability to review the diagnostic decision during follow-up allows a program to revise its erroneous or incomplete conclusions.

1.4.3 Diagnostic Style

The diagnostic style used by a program is almost as important as reaching the correct diagnosis. Although good style is hard to characterize and even harder to embody in a program, certain aspects of diagnostic style are recognizable. For example, if the program pursues some low priority diagnostic problem in the face of more important issues, if it ignores a problem of life-

threatening character, or if the stream of questions seem pointless (i.e., if the program continues to ask questions when it should have been prescribing treatment), it is likely to be rejected by the user physician.

We wish to design a program which will exhibit focused, coherent and purposeful behavior in problem solving and will know when to call a halt to its question and make an interim diagnostic judgment. In a later section we will discuss how some of these requirements can be met using notions such as *hypothetical reasoning* and *planning*.

1.4.4 Mode of Interaction

A distinction is often made between two forms of data acquisition in diagnosis: *active* and *passive* [Gorry68]. A passive mode is one in which the program is provided with all the information at one point and must make a diagnosis based on this information. An active mode is one in which the program must ask a question in order to obtain each new piece of information. The active process suffers from the shortcoming that the physician may be aware of some facts potentially useful in the diagnosis, but may not be able to communicate them to the program because each new piece of information must be requested by the program. The passive approach avoids this problem but places the responsibility of identifying relevant information on the physician. This is an unacceptable demand on a physician who is not an expert in the medical domain of the program.

Therefore, we propose a compromise position involving *mixed initiative*. In this mode, as in the active mode, the primary responsibility of gathering information still rests with the program. However, at each point in the consultation the user physician is allowed to provide a suggestion. The program must analyze this suggestion,[3] even if it chooses to ignore the suggestion as being irrelevant.

1.4.5 Handling Discrepant Information

In virtually any diagnostic workup a large amount of discrepant information must be dealt with. Some of the discrepancies arise because patients are not always accurate observers of their symptoms and because laboratory tests and medical records are often in error. In other cases a seeming discrepancy may arise because of incomplete information, i.e. there may be a valid (but so far unknown) explanation for the apparent disagreement. Correct evaluation of each type of discrepancy is critical, if the program is to perform effectively. It is necessary for a diagnostic program to be able to identify the discrepant information as it is presented in order to

---

3. The program may not, as was the case with some previous programs, put these suggestions "on hold" without reasoning about them until it is ready to ask about them. If the program does not think that the suggestion is relevant, it must make that decision explicitly.

be able to evaluate a discrepancy and choose strategies for dealing with it before incorporating it in the patient model. We have observed that the expectations of the physician play an important role in identifying possible discrepancies in the incoming information. They allows the physician to locally evaluate these discrepancies (with respect to the available evidence, physiological possibilities and the current hypothesis) and act upon them before assimilating the new information into his patient descriptions. A similar mechanism in the program is desirable. Summarizing, the importance of good handling of discrepant information can not be overstated, especially when the system is expected to be used in a normal clinical setting as well as in experimental situations.

1.4.6 Explanation

To be acceptable in an application domain such as medicine, an AIM program must go beyond providing competent advice; it must be able to explain and justify its conclusions to the user physician — much as the human consultants do today — in a language that the physician is familiar with. After all, it is the physician who provides the medical care and is primarily responsible for the welfare of the patient. It is therefore natural (even desirable) for a physician to balk at accepting advice from a "black-box" program. This reluctance perhaps accounts for much of the reported antipathy of physicians even to the programs that on statistical analysis have been shown to be as good as the expert physicians [Yu79, Kulikowski81, Long80]

We believe that a program's acceptability depends crucially upon its ability to adequately explain its reasoning and justify its conclusions. It depends on the physician being able to challenge some part of the program's conclusions and having the program explore alternatives suggested by the physician. Consultation is a "two-way street"; it can be effective only if the consultant (who is an expert in the subject matter) and the physician (who is familiar with the patient) cooperate. If any program is to be successful as an expert consultant it must allow for such an exchange.

The foregoing discussion may suggest that AIM programs be perfect, a requirement that can never be met in a real world of imperfect knowledge, where even the best of the expert physicians differ with one another. The thrust of our argument here is more limited. We are not demanding perfection from AIM programs, only that they be acceptable. Note that a program which is not as good as the best expert may nevertheless be fruitfully applied if it is acceptable and if its use improves the performance of the average clinician (who is not likely to be as good as the best expert in any given area of expertise).

In this thesis we are not extending the methodology of explanation generation. Our main thrust is in applying the available methodology to a much more complex domain than has been hitherto tried. However, since it has been demonstrated that generation of quality explanation can not be achieved by retrofitting a program with explanation capabilities, the program must be

designed with the explanation abilities in focus [Swartout80]. Our main interest is in designing explicit representation and reasoning mechanisms in the program which will provide us with the ability to justify the program's diagnoses as well as its reasoning in achieving those diagnoses.

1.5 Survey of AIM programs

Teaching of diagnostic medicine is often organized around diseases, with an emphasis on associations between the diseases and signs and symptoms typically associated with them. After all, the diagnostic task is to identify the disease hypothesis which represents the true state of the world by using all available data. Based on this observation we can conceive of a simple representation of diagnostic knowledge which draws associations between disease hypotheses and data. Given this "primitive" organization, we may already envision a diagnostic algorithm consisting of the following steps:

Diagnostic Reasoning:

(1) Whenever a new finding is reported, add it to the set of reported findings.
(2) Determine all the diseases linked to the new finding and add these diseases to the set of active hypotheses (which is initially empty).
(3) Score the active hypotheses by counting the number of expected findings observed for each disease hypothesis,
(4) rank-order the active hypotheses based on their scores and report the ranking.

Information Gathering:

(5) Select the highest-ranking disease, at least one of whose associated findings has not yet been either affirmed or denied, and ask about that finding.
(6) If step-5 fails to select a question, ask if the user is willing to volunteer a finding.
(7) If no findings are offered, report the rank ordered diagnoses and their supportive findings and stop. Otherwise, repeat steps 1 through 6.

The above algorithm, in spite of its simplicity, already captures the essential structure of a number of diagnostic programs. The association between diseases and findings forms its *static knowledge* about the domain. The set of observed findings and the rank-ordered set of active disease hypotheses are its *patient specific model* and its *understanding* of the patient's illness. The process of rank-ordering disease hypotheses is its *diagnostic evaluation*, and the selection of an appropriate finding for inquiry is its *information gathering strategy*.

The algorithm described above suffers from many inadequacies due to its oversimplification. Far more serious, however, are the problems fundamental to the model of the algorithm itself. For example, the above algorithm views diagnosis as the task of identifying that disease hypothesis which provides maximal coverage over the set of findings. Although this view of diagnosis suggests a relatively straightforward and intuitively appealing implementation, we believe this to be inadequate. Disease processes are causal; we believe that diagnosis involves providing an *adequate* explanation of the observed findings by reconstructing the possible sequence of causal events leading to the observed findings.

The program's information gathering strategy is limited to selecting one question at a time. At the end of this question, the program re-evaluates its diagnostic understanding, reformulates a new diagnostic problem (which may or may not be related to the previous problem) and selects the next question to ask. If after asking one question the diagnostic hypothesis being pursued is not confirmed, it must compete with all other active hypotheses for the attention of the diagnostic problem solver. In other words, the attention span of the program in solving any given problem is exactly one question. This results in diagnostic inefficiencies and incoherent question sequences. This problem is well recognized, and programs such as Internist-I and PIP have attempted to group diagnostic questions into meaningful packages, abating the problem somewhat. The work presented in this thesis is based on our belief that a substantial reformulation of the basic algorithm is needed before the problem can be adequately addressed [Szolovits81b].

In the remaining part of this section we will briefly review the four major AIM projects dealing with diagnosis, namely Internist-I, the Present Illness Program, CASNET/Glaucoma and Mycin. A detailed description of these programs can be found in [Szolovits81]. A good review of computer-based decision aids in medicine, using both AI and conventional computer methodologies is to be found in [Shortliffe79]. [Szolovits78] offers suggestions on the issues of choice of methodology and validation for acceptance for AIM programs. [Schwartz70] contains a discussion of acceptability issues from the viewpoint of physicians.

1.5.1 Internist-I and Present Illness Program

The Internist-I program [Pople75a, Pople77] is based on a large data base and a relatively simple evaluation and problem-selection strategy. The Internist-I data base is constructed by linking diseases and their manifestations with two subjectively assessed scores; an *evocation strength* which describes how strongly the manifestation should suggest a disease, and a *frequency* which describes how commonly the particular manifestation is observed in a patient with a given disease. Both of these are supplied by objective assessment by physicians. All the diseases are arranged into a hierarchy organized around organ-systems. Each non-terminal in this hierarchy is linked to manifestations that are common to all its inferiors. During each cycle of the algorithm, all diseases with at least one reported manifestation are evoked[4] and scored. Next, these disease hypotheses are partitioned into competing and complementary sets. This partitioning scheme represents an important contribution of the Internist-I program. It is based on two concepts: the *shelf* — a list of important manifestations that are not explained either by this diagnosis or any diagnoses previously confirmed, and the *dominance relation* — a hypothesis A is

4. If a disease (A) and one of its inferiors (B) are evoked simultaneously, then (1) if there are no known findings that can differentiate between B and any of its sibling hypotheses, B is considered to be subsumed by A and deleted from the active set. Otherwise, (2) A is replaced by the set of its immediate inferior diseases that are evoked by the manifestation.

said to dominate hypothesis B if the shelf of A is a proper subset of the shelf of B. The *competing set* is then said to contain hypotheses that either dominate or are dominated by the highest-ranking hypothesis. All other hypotheses are considered complementary and are ignored. The competing set is further reduced by considering only those hypotheses whose scores are within a fixed range of the highest-scoring hypothesis. Based on the number and relative scores of the hypotheses under consideration a diagnostic strategy (differentiate, confirm or rule-out) is selected and the next question computed. Finally, this question is asked and the diagnostic cycle is repeated.

The Present Illness Program (PIP) [Pauker76] is a frame based [Minsky75] program for taking the present illness in the domain of renal diseases. The PIP data base is implemented using disease frames, each containing the relation of the given disease to its expected findings and to other diseases, and a scoring criterion for evaluating the disease hypothesis. Some of the findings associated with a disease are specially designated as *triggers*. The complementary relation between diseases is described using *causal*, *complicational* and *associational* links; the competing relation is expressed using *differential* links. Each disease frame also contains two types of scoring functions; the *logical decision criteria* and the *numerical likelihood estimator* where the first is used for categorical evaluation and the second for probabilistic evaluation of the likelihood of the disease hypothesis under consideration [Szolovits78]. The diagnostic algorithm of PIP is similar to the basic algorithm discussed before. We should note that PIP does not use the disease-hierarchy or multiple diagnostic strategies used by the Internist-I program. On the other hand, PIP uses a substantially richer representation mechanism for describing findings and diseases as compared to Internist-I. For example, PIP allows one to describe the finding of edema observed in a given patient to be "severe", "worse in evening" and "pedal" (around legs). Finally, it uses categorical as well as probabilistic criteria for confirming diseases.

Internist-I and PIP represent medical knowledge as well as patient specific facts in phenomenological terms. The lack of physiological knowledge results in their weakness in dealing with patient illnesses with multiple interacting etiologies. The lack of physiological knowledge also results in activation of all phenomenologically possible hypotheses, including those that, based on the case-specific knowledge, are physiologically improbable. Thus, increasing the efforts needed in scoring and ruling out these hypotheses explicitly. Furthermore, the diagnostic algorithms in Internist-I and PIP alternate between obtaining a fact and evaluating the hypothesis list, resulting in a lack of focused diagnostic inquiry as discussed before.

The patient-specific model in Internist-I and PIP consists of a collection of patient facts and the list of active hypotheses; it does not relate different findings and hypotheses into causal explanations. As a result these programs have only a fragmentary understanding about the patient's condition and they often change their description of the patient's illness radically without substantial indications to that effect.

1.5.2 CASNET/Glaucoma

The Glaucoma program deals with the diagnosis and treatment of eye diseases. It is implemented using the CASNET [Weiss74] theory of representation of causal knowledge. The medical knowledge in Glaucoma is represented as a network of physiological states. These states are linked together by subjectively assessed transition probabilities, and by support values indicating how strongly certain test results support the presence of a particular condition (state). The transitional probabilities are used primarily as a means of selecting the most appropriate next state to investigate and the support values are used to evaluate the score (fuzzy likelihood [Gaines76, Zadeh65]) of a state, which is used to confirm or deny a state. Finally, the patterns of confirmed and denied states in the network are interpreted using a number of programs which compare the progress of the diseases in the given patient with the diseases known to the individual program.

The use of physiological knowledge gives the glaucoma program a better understanding of the mechanisms of disease evolution and interaction than the other programs discussed above. However, its use of causal knowledge is restricted to the local propagation of likelihood weights to determine the most appropriate next state for investigation. The program cannot use hypothesized diagnoses to guide its diagnostic inquiry: it separates the process of information gathering from that of diagnosis. The information gathering is directed solely towards confirming (or ruling out) states in the causal net.[5] Moreover, the program works in a domain where the disease physiology is uniformly well understood and each state can be confirmed directly using some test. Therefore, the techniques developed in this program are not easily extendable to programs working in other domains of medical expertise.

1.5.3 Mycin

Mycin is a rule-based program [Shortliffe76, Davis77] for diagnosis and treatment of infectious diseases — in particular, bacterial infections in the blood (and recently extended to other infectious diseases). It represents medical knowledge in terms of production rules [Davis77] and uses a collection of associative triples to represent the patient specific knowledge [Shortliffe75, Shortliffe76]. A novel mathematical model of confirmation [Shortliffe76] selects a set of organisms suspected of causing the illness. Diagnosis is carried out using a simple goal-directed control structure with backward chaining. The highest-level goal of Mycin is to determine if the patient is suffering from a significant infection which should be treated, and if he is, to select the appropriate therapy. It retrieves all the rules applicable to this goal and applies them sequentially as follows. It attempts to ascertain whether the "conclusion" of a rule is valid by evaluating each of its premises. If this information is already available in the data base, the

---

5. During this phase the program does not attempt to identify diseases responsible for the presence of these states. The diagnosis is attempted separately after the information gathering phase is completed.

program retrieves it. If not, determination of this premise becomes the new goal, and the program recurs. If after trying all the relevant rules, the answer still has not been discovered, the program asks the user for the relevant clinical information which will permit it to establish the validity of the premise clause. Thus, the rules "unwind" to produce a succession of goals, and it is this attempt to achieve each goal that drives the consultation.

The rules in Mycin are used to represent the domain knowledge as well as to encode the flow of control of the program. This takes away some of the advantages of modularity of knowledge because one must take into account the possible interactions between rules during problem solving. The goal structure of Mycin allows efficient problem solving and can be used for explaining the problem solving behavior of the program, but the program cannot explain the medical significance of its behavior as this information is compiled out while writing the rules.

The rule-based Mycin methodology is applicable in fields where the domain specific knowledge can be described using judgmental rules. It appears to require a field which has attained a certain level of formalization with a generally recognized set of primitives and a minimal understanding of basic processes and which does not have a high level of interaction between conceptual primitives [Davis77]. Finally, the rule-based methodology developed by Mycin and its derivative programs can be used effectively in encoding knowledge needed in handling specific well defined situations such as special heuristics for differentiation between two similar diseases which are difficult to differentiate using global differentiation heuristics.

The programs described above can be classified as the "first generation AIM programs". These programs have contributed immensely by demonstrating the feasibility of using computers (and AI techniques) in medical diagnosis. Some of the significant developments in this regard are summarized here.

The active hypothesis set introduced in PIP and the hierarchic organization of diseases introduced in Internist-I provide useful techniques for organizing programs for efficiency. A heuristic to partition the hypothesis set into competing and complementary sets was introduced in Internist-I. In spite of its shortcomings, the partitioning heuristic is intuitively appealing and empirically effective [Pople75a]. An improved technique for identifying complementary and competing hypotheses, especially for illnesses caused by multiple diseases, is one of the topics of interest in this thesis.

Recognizing that pathognomonic and important evocative findings help to focus the diagnostician's attention sharply, mechanisms to flag such findings and their use in focusing the programs attention were developed in Internist-I and PIP. Heuristics to help confirm or eliminate hypotheses categorically (without resorting to revised probabilities and thresholds) and explicit differential diagnosis links to indicate well-known points of diagnostic confusion were also added in PIP.

Causality as a major mechanism for tying together independent hypothesized disorders was identified as a fundamental mechanism in the CASNET/Glaucoma program, Internist-I and PIP. The Glaucoma program went a step beyond the others in the use of causality by defining disease as a progression of causally connected states. However, in all three programs, the use of causality is limited to propagating probability-like estimates of likelihood which remain the primary criterion for their clinical decisions.

The need for explanation and justification capabilities in an AIM programs was first recognized by and implemented in MYCIN. In this chapter we have argued that these capabilities are essential for the success of any consulting program. In this thesis we take this capability to be an essential component of the design of ABEL program.

1.6 Outline of the Thesis

This thesis contains seven chapters and two appendices. Chapter 2 previews the capabilities of the program with the help of two simple examples. Chapter 3 describes the representation of ABEL's medical knowledge. The medical knowledge consists of a hierarchic representation of anatomical, physiological, etiological and temporal knowledge. This forms the groundwork for an efficient representation of diseases and their pathophysiology in the domain of electrolyte and acid-base disturbances. The diseases are defined in terms of their loci along these four dimensions, providing a natural hierarchic organization to the disease definitions. This framework of basic medical knowledge provides us with a vocabulary for expressing phenomenological and pathophysiological knowledge.

An expert physician may have an understanding of a difficult case in terms of several levels of detail. As noted earlier, at the shallowest level that understanding may be in terms of commonly occurring associations of syndromes and diseases, whereas at the deepest it may include the biochemical and pathophysiological interaction of abnormal findings. Chapter 3 describes a multi-level description of pathophysiology, where each level of description can be viewed as a semantic net of relations between diseases and findings. Each node in the net represents a normal or abnormal state and each link represents a relation (causal, associational, etc.) between these states. Each node is associated with a set of attributes describing the temporal characteristics, severity or value, and other relevant attributes. Each link describes a causal relation between a cause node and an effect node by specifying a multivariate relation between attributes of the cause and the effect. Additional information to support mapping knowledge at one level to an adjacent level is also described.

In Chapter 4, we propose the use of a *coherent hypothesis* as the logical unit of hypothesis representation. This captures our notion, expressed above, that the reasoner's hypothesis structure must account for the total state of mind of the reasoner including its current uncertainties. In the program, each coherent hypothesis is represented using a *patient specific*

model (PSM). Each PSM represents a causal explanation of all the observed findings and their interrelationships at various levels of detail. Note that within each PSM all the diseases, findings, etc., are mutually complementary, while the alternate PSM's are mutually exclusive and competing.

The PSM is created by instantiating portions of ABEL's general medical knowledge and filling in its details from the specific case being considered. The instantiation of the PSM is very strongly guided by initially given data, because the PSM includes only those disorders and connections that are needed to explain the current case. Instantiation is accomplished by five major operators. *Initial formulation* creates an initial patient description from the presenting complaints and laboratory results. *Aggregation* and *elaboration* make connections between the levels of detail in the PSM by filling in the structure above and below a selected part of the network, respectively. In a domain such as ABEL's, multiple disorders in a single patient and the presence of homeostatic mechanisms require the program to reason about the joint effects of several mechanisms which collectively influence a single quantity or state. *Component decomposition* and *summation* relate disorders at the same level of detail by mutually constraining a total phenomenon and its components; the net change in any quantity must be consistent with the sum of individual changes in its parts. The final operator, *projection*, forges the causal links within a single level of detail in the search for causal explanations. The operators all interact because the complete PSM must be self-consistent both within each level and across all its levels. Therefore, each operation typically requires the invocation of others to complete or verify the creation of related parts of the PSM. Furthermore, PSM's are organized in a context tree allowing different PSM's to share structures common to them. The root of the PSM-tree also contains all the observed findings and diseases which have been concluded to be true so that they may be shared by all PSM's.

Locality is a desirable property for the reasoning and description schemes. It imposes modularity in the organization of knowledge, making acquisition and representation of knowledge tractable. Furthermore, it makes possible efficient reasoning schemes whose resource requirements do not grow with increasing size of the data-base.[6] To exploit the locality constraint in reasoning with causal networks, a program should be able to reason based only on the information locally available from the neighborhood of the mechanism under consideration. Although it is always possible to choose a level of abstraction at which the interaction between a given pair of states can be described locally, for a given level of detail it is not possible to impose the locality constraint on every interaction. The multiple-level causal model and the abstraction/elaboration process presented in this thesis allow us to overcome this problem. For example, if at some level of detail two distant states interact, we can aggregate the description of

---

6. Locality has been exploited in a large number of diverse problems, such as common-sense reasoning [Minsky73, Kuipers77, deKleer79] and natural language processing [Marcus79, Church80, Martin81]. For example, the constraint of "context freeness" in natural language is a specific instance of locality constraint.

intervening causal network to a level where the two states are adjacent to one another. The interaction between the two can now be computed locally.

Chapter 5 discusses the diagnostic problem solving activity. The diagnostic problems are formulated by identifying the weaknesses and conflicts in the PSM's. The task of the diagnostic problem solver is to resolve these conflicts and weaknesses by gathering new information. We note that the medical knowledge in the program consists of prototypes of the disease entities. However, this prototypical knowledge can be substantially constrained because the hypothesized disease entities must be consistent with the known facts and explanations. We introduce the notion of a *diagnostic closure* which extracts and tailors that part of medical knowledge that is directly relevant to the diagnostic task at hand. The diagnostic closure brings together all the dependencies and expectations necessary for planning a diagnostic inquiry, for evaluating real and apparent discrepancies in the incoming information, and provides a framework for explaining the alternatives under consideration and for justifying the selection of questions. Although we envision using recent advances in the planning paradigm [Fikes72, Sacerdoti75, Stefik81], the current implementation of the program generates a simple tree-structured plan for information gathering by decomposing the problem by successive applications of confirm, rule-out, differentiate, and group-and-differentiate strategies. Finally, when a sufficient amount of new information is available the program assimilates this information into the PSMs and the diagnostic process is repeated. The process terminates when an adequate explanation for the patient's illness is found or when all the information necessary for such an explanation is exhausted.

In chapter 6 we revisit the example described in chapter 2 in greater detail. Chapter 7 summarizes the experience gained and lessons learned in this enterprise and indicates pointers to future research. Finally, appendix 1 briefly summarizes the XLMS system (a knowledge representation system built on top of LISP) used by ABEL. Appendix 2 summarizes the techniques for translating the internal data structures of the program into English developed recently by Swartout [Swartout80] and discusses algorithms for organizing the concepts encoded in causal networks into a linear sequence of sentence level objects that can then be translated using the above-mentioned methodology.

2. Examples

This chapter presents the inner workings of ABEL with the help of annotated examples. In this chapter the reader is not expected to understand *how* the program accomplishes its task, but rather just *what* it does. The succeeding chapters will examine the structure of the program and the method by which each step is accomplished. We will consider two examples: (1) a patient suffering from moderately severe salmonellosis, and (2) a patient suffering from moderately severe salmonellosis and vomiting. The selection of the medical examples is motivated by our desire to make the medical contents of the examples as simple as possible. In chapter 6 we will revisit these examples and discuss how the program accomplishes each of its tasks.

2.1 Example 1: Salmonellosis

For the first example let us consider a 40 year old 70 Kg male patient who has been suffering from moderately severe salmonellosis and, as a result, has developed moderately severe metabolic acidosis and hypokalemia. To illustrate the program let us provide it initially with only the laboratory analysis of the patient's blood sample (serum analysis) without any clinical information.

```
Serum Analysis:
    Time:   0                    time of the session
     Sex:   male
      Na:   142 meq/l            normal
       K:   3   meq/l            moderately low
      Cl:   113 meq/l            normal
    HCO3:   15  meq/l            moderately low
    pCO2:   30  mmHg             moderately low
```

Based on these data, the program generates all possible acid-base disturbances that can account for the laboratory data. It then prunes and rank-orders these disturbances based on their complexity, likelihood and severity of each component. The rank-ordered list of likely disturbances is:

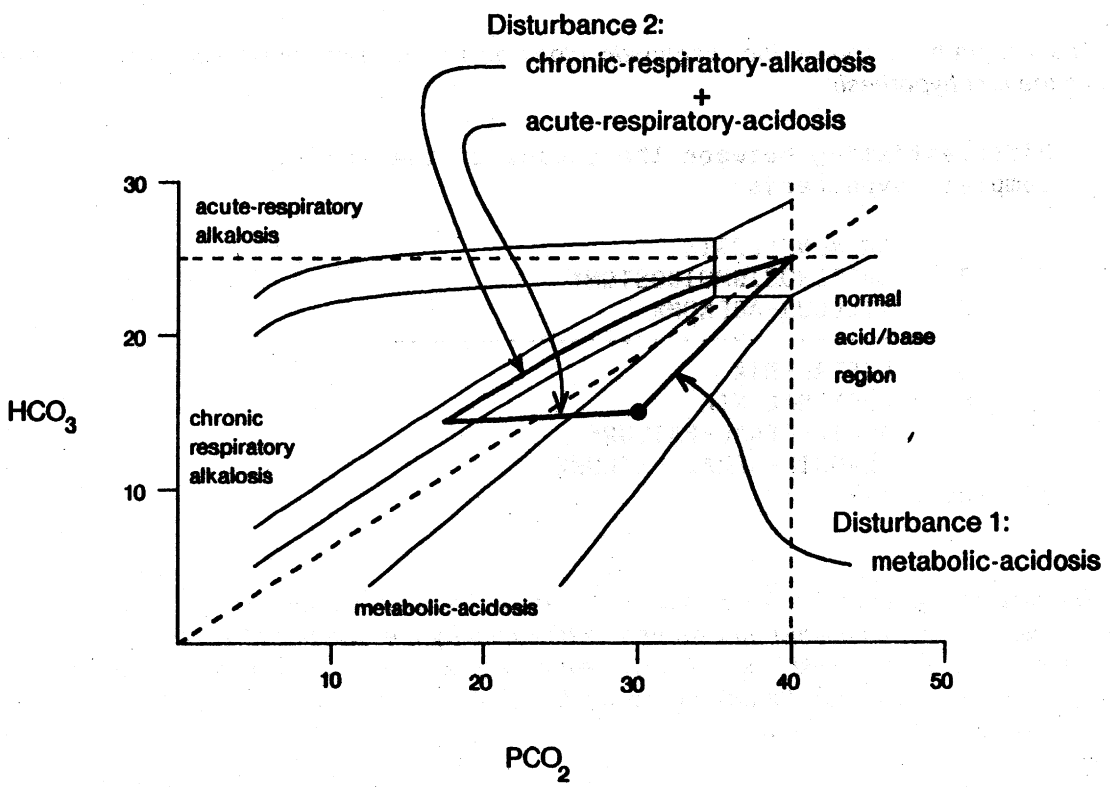
```
---- Patient Acid-Base Profile ----

1. metabolic-acidosis [severity: 0.4]                    very likely
2. chronic-respiratory-alkalosis [severity: 0.68]
   + acute-respiratory-acidosis [severity: 0.32]         unlikely
```

The computation of the acid-base profile is based on the Nomogram of Acid-Base Disturbances described in chapter 1. Figure 4 shows the relevant region of this nomogram with the loci of the two hypothesized disturbances. The estimation of the severity of a disturbance is

based on the length of the segment along the locus of that disturbance. Thus, we note that the severity of the single acid-base disturbance (metabolic acidosis) is only about 0.4 while an equivalent acid-base disturbance composed of chronic respiratory alkalosis and acute respiratory acidosis has severities of 0.68 and 0.32 respectively.[7]

Next, the program creates a PSM for each possible acid-base disturbance and interprets the laboratory data in the context defined by each acid-base disturbance. For example, with the assumption of fully compensated metabolic acidosis, the entire change in the PCO_2 may be considered chronic, therefore, the chronic component of the PCO_2 will be 30 meq/l, while with the assumption of chronic respiratory alkalosis and acute respiratory acidosis, the chronic component of PCO_2 is due only to the chronic component of this disturbance, therefore reading from the nomogram we find that the chronic value of PCO_2 in this case will be approximately 16 meq/l.

Fig. 4. Graphic depiction of the two Acid-Base hypotheses



7. The numbers corresponding to the acid-base disturbances computed above are the programs internal assessment of the severity of illness, they are not measurable.

The program then aggregates its patient-specific physiologic knowledge to formulate an interpretation of the laboratory data at the clinical level. The computer generated explanation of its interpretation of these data under the two major hypotheses is described in figures 5 and 6.

A quick look at the two clinical level explanations shows that the structure involving hypokalemia and acidemia is common to the two hypotheses. They differ in their accounting for acidemia. Note that the clinical level abstraction of the two hypotheses is fairly simple in structure and does not contain any feedback cycles. The cycles present at the intermediate level describing the interaction between the acidemia, hypobicarbonatemia and hypocapnia have been abstracted away. A closer look at these feedback cycles shows the principal difference between the two hypotheses. In the first case, the change in the acid-base state is a consequence of loss of HCO_3 from the body which causes hypobicarbonatemia, whereas in the second it enters as primary disturbance in ventilation which alters the PCO_2. Finally, we note that the first hypothesis has two unaccounted findings while the second hypothesis contains three unaccounted findings.

In the context of this initial analysis of the patient's condition, the program starts the diagnostic exploration. An annotated (in italics) transcript of the program's diagnostic behavior is shown next.

The program computes the diagnostic closures for the two hypotheses and decides to pursue the first hypothesis.

```
Differentiating between the causes of the leading
complete hypothesis.

     1    SALMONELLOSIS
     2    URETEROSIGMOIDOSTOMY
     3    VILLOUS-ADENOMA
     ------------------------------------
     4    DISTAL-RTA
     5    PROXIMAL-RTA
     6    ACUTE-RENAL-FAILURE
     7    CHRONIC-RENAL-FAILURE
continue? ==>
```

The list above contains all possible diseases that can explain some part of the first hypothesis. The list is divided into groups of diseases by the number of unaccounted findings that each disease can explain succinctly. Within each group the diseases are ordered by a secondary scoring criterion based on the quality of their match with the hypothesis and their potential to be ultimately confirmed.

```
Differentiating between
     SALMONELLOSIS URETEROSIGMOIDOSTOMY VILLOUS-ADENOMA
```

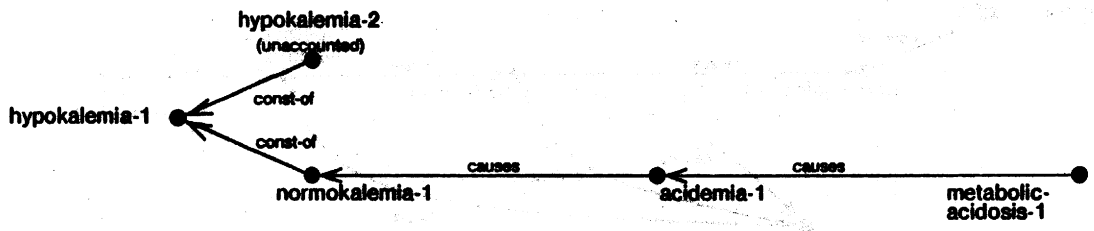
Fig. 5. Comparision of hypotheses 1 & 2 at clinical level

Hypothesis 1: Metabolic Acidosis

This is a 40 year old 70.0 kg male patient. His electrolytes are:

Na: 142.0	HCO3: 15.0	Anion Gap: 13.0
K: 3.0	pCO2: 30.0	
Cl: 113.0	pH: 7.32	

The patient has moderate metabolic acidosis and mild hypokalemia. The metabolic acidosis causes mild acidemia. The acidemia partly compensates the suspected moderate hypokalemia leading to the observed hypokalemia. The metabolic acidosis remains to be accounted for. The hypokalemia has only been partially accounted for.



Hypothesis 2: Chronic Resp. Alkalosis & Acute Resp. Acidosis

This is a 40 year old 70.0 kg male patient. His electrolytes are: ...

The patient has moderate chronic respiratory alkalosis, moderate acute respiratory acidosis and mild hypokalemia. The acute respiratory acidosis and chronic respiratory alkalosis cause mild acidemia. The acidemia partly compensates the suspected moderate hypokalemia leading to the observed hypokalemia. The chronic respiratory alkalosis and acute respiratory acidosis remain to be accounted for. The hypokalemia has only been partially accounted for.

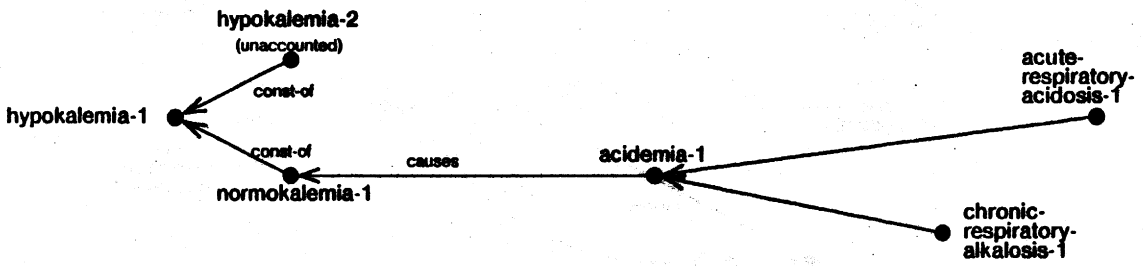
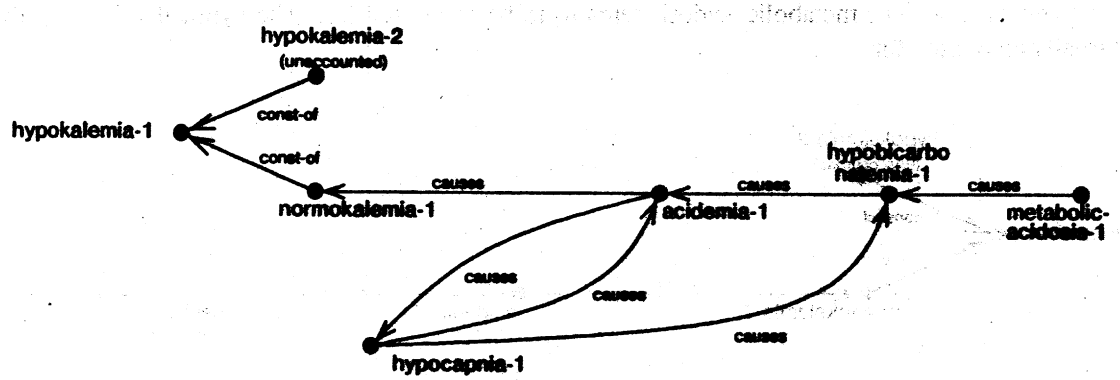


Fig. 6. Comparision of hypotheses 1 & 2 at intermediate level

Hypothesis 1: Metabolic Acidosis

This is a 40 year old 70.0 kg male patient. His electrolytes are: ....

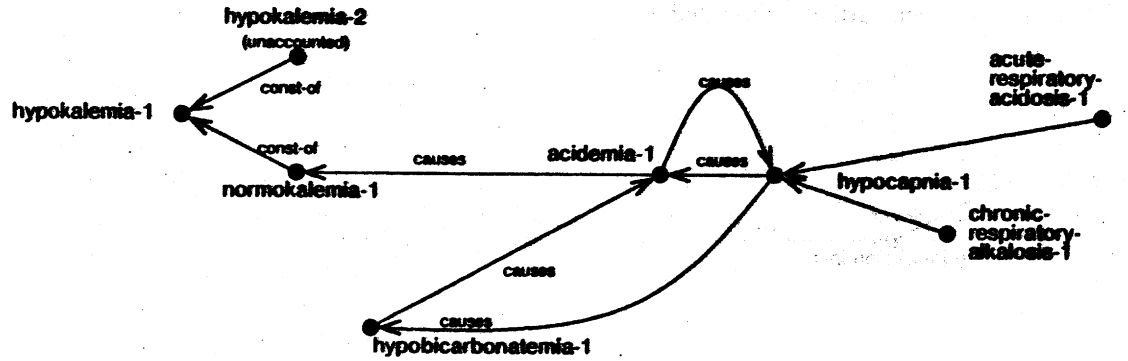
The patient has moderate metabolic acidosis, mild hypokalemia and moderate hypobicarbonatemia. The metabolic acidosis along with moderate hypocapnia causes hypobicarbonatemia. The hypobicarbonatemia along with hypocapnia causes mild acidemia. The acidemia partly compensates the suspected moderate hypokalemia leading to the observed hypokalemia. The metabolic acidosis remains to be accounted for. The hypokalemia has only been partially accounted for.



Hypothesis 2: Chronic Resp. Alkalosis & Acute Resp. Acidosis

This is a 40 year old 70.0 kg male patient. His electrolytes are: .....

The patient has moderate acute respiratory acidosis, moderate chronic respiratory alkalosis, mild hypokalemia and moderate hypobicarbonatemia. The chronic respiratory alkalosis and acute respiratory acidosis along with mild acidemia cause moderate hypocapnia, which causes hypobicarbonatemia. The hypobicarbonatemia and hypocapnia cause acidemia. The acidemia partly compensates the suspected moderate hypokalemia leading to the observed hypokalemia. The acute respiratory acidosis and chronic respiratory alkalosis remain to be accounted for. The hypokalemia has only been partially accounted for.



As a first step towards differentiation, the program asks if the user is already aware of any of the possible alternatives.

```
Does the patient have any of the following?
     1    SALMONELLOSIS
     2    URETEROSIGMOIDOSTOMY
     3    VILLOUS-ADENOMA

Please enter numbers between 1 and 3
Present: ==> E(xplain)
```

The user indicates that he would like an explanation. The program prompts by indicating in a menu the possible choices for explanation.

```
     1    DESCRIBE-HYPOTHESIS
     2    DESCRIBE-DIAGNOSTIC-REASONING
     3    DESCRIBE-FINDING
     4    DESCRIBE-CAUSE
     5    DONE
==> 2
The stack of diagnostic goals:

     3    DIFF-CAUSES
     2    CONFIRM-CH
     1    DIAGNOSE

    I am diagnosing the possible acid base disturbances.  I am
confirming the leading hypothesis.  I am differentiating between
its leading causes.

     1    DESCRIBE-HYPOTHESIS
     2    DESCRIBE-DIAGNOSTIC-REASONING
     3    DESCRIBE-FINDING
     4    DESCRIBE-CAUSE
     5    DONE
==> 4

Please enter a cause for viewing: ==> salmonellosis

    Salmonellosis may account for moderate metabolic acidosis and
moderate hypokalemia as follows.
    Moderate salmonellosis may cause metabolic acidosis.
    Moderate salmonellosis may cause hypokalemia

     1    DESCRIBE-HYPOTHESIS
     2    DESCRIBE-DIAGNOSTIC-REASONING
     3    DESCRIBE-FINDING
     4    DESCRIBE-CAUSE
     5    DONE
==> 5
```

```
Does the patient have any of the following?
     1    SALMONELLOSIS
     2    URETEROSIGMOIDOSTOMY
     3    VILLOUS-ADENOMA

Present: ==> none      Absent: ==> none      Unknown: 1 2 3

Differentiating between
     SALMONELLOSIS URETEROSIGMOIDOSTOMY VILLOUS-ADENOMA
```

Salmonellosis, ureterosigmoidostomy and villous-adenoma all cause a reduction in extracellular fluid, resulting in dehydration. However, the program notices that some of the diseases in the second set (e.g., renal failure) may have the exact opposite effect of causing edema. Therefore, while exploring the state of extracellular fluid the program includes edema in the question.

```
Does the patient have one of the following?
     1    DEHYDRATION
     2    EDEMA

Present: ==> none      Absent: ==> none      Unknown: 1 2
```

The program is expecting dehydration. Therefore, when we fail to confirm or deny the dehydration the program pursues the finding further.

```
I would like to ask about the effects of SALMONELLOSIS.

Is the value of SERUM-CREATININE known?      ==> E(xplain)


    High serum creatinine may be caused by moderate salmonellosis
as follows:

     Moderately high serum creatinine may be caused by moderate
dehydration, which may be caused by salmonellosis.
```

The user indicates that he would like a justification for this question. The program generates the explanation by tracing back the causal path from serum creatinine in the DC associated with the goal of the question.

The program's diagnostic reasoning at this point can be described as follows: The top level goal of the program is to do diagnosis. In order to do the diagnosis the program evaluates the two alternate hypotheses and selects the first hypothesis (metabolic-acidosis) for confirmation. To confirm this hypothesis it selects the set of leading causes for the first hypothesis, namely, salmonellosis, ureterosigmoidostomy and villous-adenoma. It then attempts to differentiate between these causes. It determines the findings predicted by each of these three causes, orders them according to their discriminatory power, and asks about them, the most discriminating finding first.

```
Is the value of SERUM-CREATININE known? ==> yes

Please enter the attributes of SERUM-CREATININE
What is the VALUE of SERUM-CREATININE ? ==> 3
What is the START-TIME of SERUM-CREATININE ? ==> 0

Is the value of MEAN-ARTERIAL-BLOOD-PRESSURE known? ==> yes

Please enter the attributes of MEAN-ARTERIAL-BLOOD-PRESSURE
What is the VALUE of MEAN-ARTERIAL-BLOOD-PRESSURE ? ==> 75
```

The program has now completed one full cycle of its planned diagnostic inquiry. It now incorporates this information into both hypotheses and starts the next cycle of diagnostic planning.

```
Starting next cycle of diagnosis

Differentiating between the causes of the leading complete hypothesis.

     1    SALMONELLOSIS
     ------------------------------------
     2    VILLOUS-ADENOMA
     3    URETEROSIGMOIDOSTOMY
     ------------------------------------
     ------------------------------------
     4    ADRENAL-INSUFFICIENCY
     5    DIABETES-INSIPIDUS
     6    ACUTE-RENAL-FAILURE
     7    CHRONIC-RENAL-FAILURE
     ------------------------------------
     8    DISTAL-RTA
     9    PROXIMAL-RTA
continue? ==>
```

The program has already gathered sufficient information to confirm salmonellosis. It is unable to do so because we have not implemented the criteria for confirming a disease yet. However, we note that the information gathered has provided substantial categorical separation between the gastrointestinal and the renal causes of the metabolic-acidosis with hypokalemia.

```
Pursuing SALMONELLOSIS: no findings available

Differentiating between VILLOUS-ADENOMA URETEROSIGMOIDOSTOMY:
     no finding available

Differentiating between ADRENAL-INSUFFICIENCY DIABETES-INSIPIDUS
ACUTE-RENAL-FAILURE CHRONIC-RENAL-FAILURE
```

As there are more than two items in the differentiation set, the program groups these items into renal and extra-renal sets. It then pursues the renal set first.

```
Does the patient have any of the following?
     1    CHRONIC-RENAL-FAILURE
     2    ACUTE-RENAL-FAILURE

Present: ==> none      Absent: ==> 1 2

Does the patient have any of the following?
     1    ADRENAL-INSUFFICIENCY
     2    DIABETES-INSIPIDUS

Present: ==> none      Absent: ==> 1      Unknown: 2

Differentiating between DISTAL-RTA PROXIMAL-RTA

Does the patient have any of the following?
     1    PROXIMAL-RTA
     2    DISTAL-RTA

Present: ==> none      Absent: ==> none      Unknown: 1 2
```

The program has now completed the second cycle of its planned diagnostic inquiry. No new finding was uncovered during this cycle of diagnosis and therefore, the program does not need to revise the two hypotheses. However, during the planning of the next cycle of diagnostic inquiry it discovers that all the questions relevant to the etiology of metabolic-acidosis and hypokalemia have already been exhausted. Therefore, in order to proceed it requests that some possible etiology be assumed.

```
Starting next cycle of diagnosis
No new finding added in the previous cycle.

    All possible etiologies that could explain the patient's
illness are unknown.  In order to proceed we must at least
hypothetically assume one of them.  Possible etiologies that could
explain the patient's illness listed in decreasing order are:

     1    SALMONELLOSIS
     ------------------------------------
     2    VILLOUS-ADENOMA
     3    URETEROSIGMOIDOSTOMY
     ------------------------------------
     ------------------------------------
     4    DIABETES-INSIPIDUS
     ------------------------------------
     5    DISTAL-RTA
     6    PROXIMAL-RTA

Would you like to assume SALMONELLOSIS ?      ==> yes

Assuming MODERATE ACUTE SALMONELLOSIS.
```

The program adds salmonellosis to the patient models and re-evaluates two hypotheses before resuming the diagnosis again. However, it immediately recognizes that it has found the missing piece in the puzzle! Based on the assumption that the patient has salmonellosis, the program selects the first PSM (metabolic acidosis) as the explanation of the patient's acid-base disorder. The program's explanations of the successful and the alternate PSMs are shown next. Note, however, that these explanations are provided with the assumption that the patient has salmonellosis, an assumption that needs to be verified.

```
Diagnosis completed.  The successful diagnosis is:

    This is a 40 year old 70.0 kg male patient with moderate
salmonellosis.  His electrolytes are:

        Na: 142.0          HCO3: 15.0        Anion Gap: 13.0
         K: 3.0            pCO2: 30.0
        Cl: 113.0           pH: 7.32       Creatinine: 3.0

    The salmonellosis causes moderate metabolic acidosis and
moderate dehydration.  The dehydration causes moderate
hypotension and moderate high creatinine disturbance.  The
metabolic acidosis causes mild acidemia.  The salmonellosis and
acidemia cause mild hypokalemia.  All findings have been
accounted for.
```

As can be seen from the above explanation, the program has concluded that salmonellosis provides an adequate explanation for the patient's illness and that the first hypothesis is substantially superior to the alternate hypothesis. The alternate hypothesis with all the available information added is shown below.

```
    This is a 40 year old 70.0 kg male patient with salmonellosis.
His electrolytes are:

        Na: 142.0          HCO3: 15.0        Anion Gap: 13.0
         K: 3.0            pCO2: 30.0
        Cl: 113.0           pH: 7.32       Creatinine: 3.0

    The salmonellosis causes moderate metabolic acidosis and
moderate dehydration.  The dehydration causes moderate
hypotension and moderate high creatinine disturbance.
Moderate acute respiratory acidosis, moderate chronic
respiratory alkalosis and metabolic acidosis partly
compensate the suspected mild alkalemia leading to the
observed mild acidemia.  The salmonellosis and acidemia cause
mild hypokalemia.  The chronic respiratory alkalosis and
acute respiratory acidosis remain to be accounted for.  The
alkalemia has only been partially accounted for.
```

Notice the difference in the two explanations. The first explanation contains only one acid-base disturbance, while the second explanation contains a total of three acid-base disturbances. Furthermore, all the findings in the first hypothesis have been accounted for while the second hypothesis has three acid-base disturbances still to be accounted for.

2.2 Example 2: Vomiting and Salmonellosis

The next example illustrates the program's capabilities in dealing with multiple etiologies, namely salmonellosis and vomiting, which offset the effects of each other on the acid-base balance. We will focus on the program's understanding and its ability to reformulate this understanding when new information is provided.

Let us consider a patient who is suffering from moderately severe vomiting for the past two days who then develops salmonellosis. Note that the electrolyte and acid-base disturbances in vomiting result from the excessive loss of upper gastrointestinal fluid, while in salmonellosis they result from the loss of lower gastrointestinal fluid. The upper GI fluid is acidic while the lower GI fluid is alkaline, therefore the two tend to have offsetting effects on the acid-base balance. However, vomiting and salmonellosis both cause hypokalemia and dehydration, therefore they compound these effects of each other. For this example, let us consider a patient in which the presentation of vomiting and salmonellosis are such that each exactly cancels the acid-base effect of the other, leaving the patient with no acid-base disturbance. We will illustrate the program's handling of this case by describing the program's understanding of the case at three points during the diagnostic process: (1) just after the electrolyte values are entered in the program, (2) after the finding of vomiting has been presented, and (3) at the end of the diagnostic process.

The program's evaluation of the serum electrolytes and the English explanation of its initial hypothesis are:

```
Serum Analysis:
    Time:  0
     Sex:  male
      Na:  141  meq/l       normal
       K:  2    meq/l       low
      Cl:  108  meq/l       normal
    HCO3:  25   meq/l       normal
    pCO2:  39   mmHg        normal

    ---- Patient Acid-Base Profile ----
    1.  normal-acid-base-state

    This is a 40 year old 70.0 kg male patient with moderate
hypokalemia.  His electrolytes are: .....
```

The serum analysis reveals only one abnormal finding, hypokalemia. The program starts the diagnostic process by attempting to differentiate between the possible causes of hypokalemia which include vomiting and salmonellosis along with other etiologies such as laxative abuse, diuretic use, hyperaldosteronism etc. The summary of the program's hypothesis after the finding

of moderately severe vomiting has been presented is:

```
    This is a 40 year old 70.0 kg male patient with moderate
vomiting.  His electrolytes are:

      Na: 143.0          HCO3: 25.0        Anion Gap: 12.0
       K: 2.0            pCO2: 39.0
      Cl: 108.0            pH: 7.42

    The vomiting causes moderate metabolic alkalosis.  Moderate
hypokalemia is partly caused by vomiting leaving some
additional factor causing hypokalemia still unaccounted for. The
hypokalemia and moderate acidemia have only been partially
accounted for.
```

Notice that the vomiting partially accounts for the observed hypokalemia. However, in order to account for the hypokalemia the program must assume that there has been substantial upper GI fluid loss sufficient to also cause metabolic alkalosis. As this metabolic alkalosis is not consistent with the normal acid-base state the program must decompose the normal acid-base state into offsetting alkalemia and acidemia. The alkalemia which is accounted for by metabolic alkalosis, and acidemia which remains unaccounted for. The remaining unaccounted components now present a picture similar to that of the previous case (example 1) and the diagnosis proceeds similarly. The diagnosis is completed when the program is told about salmonellosis (the remaining disturbance). A summary of the programs' final diagnosis is described next.

```
    This is a 40 year old 70.0 kg male patient with moderate
vomiting and moderate salmonellosis.  His electrolytes are:

      Na: 143.0          HCO3: 25.0        Anion Gap: 12.0
       K: 2.0            pCO2: 39.0
      Cl: 108.0            pH: 7.42       Greatinine: 3.0

    The vomiting causes moderate metabolic alkalosis.  The
salmonellosis and vomiting cause moderate dehydration, which
causes moderate hypotension.  The dehydration also causes
moderate high creatinine disturbance.  The salmonellosis causes
moderate metabolic acidosis.  The metabolic acidosis and
metabolic alkalosis cause normal ph.  The salmonellosis, normal
ph and vomiting cause moderate acute hypokalemia.  All findings
have been accounted for.
```

The primary focus of this thesis is in developing a methodology for knowledge representation and manipulation that allows our program to exhibit the understanding of patient illness demonstrated above. In the next three chapters we will study in detail this methodology and its

implementation before revisiting the same examples again in greater detail.

3. Representation of Medical Knowledge

Illness can be described as a change in the normal state or function in a patient. To describe an illness, we need a formalism to represent the states, the state changes, the normal and the abnormal functions and their interactions in terms of the primitives known to the system. This knowledge is organized in the program with the help of (1) an anatomy component, which includes a *part-of* hierarchy for organ systems, *contained-in* and *position* relations for major anatomical features, and a *connected-to* relation which provides material flow information. (2) A physiology component, where our concentration has been only on the fluid and electrolytes, describes the fluid compartments of the body, the spaces of distribution of various solutes, and the relative distribution of losses and gains in the various compartments under different conditions. (3) A pathophysiology component, which contains some primitive knowledge about disease etiologies, a taxonomy of disease processes, and causal relations which describe how the changes in a given state influence other states.

It is also important to recognize commonly occurring constellations of abnormal states as special composite situations. Conceptualization of these composite situations in a diagnostic system is important because it provides us with the ability to reason at a high level of abstraction, and to organize a large number of seemingly unrelated facts into a coherent whole. We have argued that it is crucial for any diagnostic system to have the ability to reason simultaneously at a high level of abstraction consisting of phenomenological knowledge as well as at a physiological level. We accomplish this with the help of a multi-level model for representation of diseases and causal phenomena. This is motivated by the observations made by Lynch while studying the *conceptual maps* of metropolitan regions. He notes

> *"Rather than a single comprehensive image for the entire environment, there seemed to be sets of images, which more or less overlapped and interrelated. They were typically arranged in a series of levels, roughly by the scale of area involved, so that the observer moved as necessary from an image at street level to levels of a neighborhood, a city, or a metropolitan region."*
>
> — The image of the city [Lynch60, pages 85-86].

The structure of the cognitive map described above is a product of the necessity to cope with *large-scale maps*; maps that are too large to be perceived at once, too large to be stored in the short-term memory by their users at a single instance of time, and too complex to be computationally tractable in solving problems (such as finding an efficient path between two points on the map). An important observation in formulating cognitive maps is that they are organized around *landmarks*. The conceptualization can be achieved by expanding the denotation of a landmark to subsume the local topology surrounding the designated location. If this conceptualization is carried out carefully, so that the areas subsumed by these landmarks overlap and cover the entire detailed map, it is possible to maintain sufficient coherence

(mapping) to be able to move between different levels of description.

Based on these observations and similar observations of a physician's use of medical knowledge, we have developed a hierarchical multi-level representation scheme to describe medical knowledge. The lowest level of this description consists of pathophysiological knowledge about diseases, which is successively aggregated (summarized) into higher level concepts and relations, gradually shifting the content of the description from physiological to syndromic knowledge. The aggregate syndromic knowledge provides us with a concise global perspective and helps in the efficient exploration of the diagnostic alternatives. The physiological knowledge, on the other hand, provides us the capabilities of handling complex clinical situations arising in patients with multiple disturbances, evaluating the physiological validity of the diagnostic possibilities being explored, and organizing a number of fragmented and seemingly unrelated facts into a coherent causal description.

3.1 Anatomical Knowledge

The anatomical knowledge of the system includes (1) a *part-of* hierarchy for organ systems, (2) *connected-to* relations, which provide the material flow information, and (3) *contained-in* and *position* relations which provide gross anatomical relations between anatomical entities.

3.1.1 Anatomical Taxonomy

The part-of hierarchy defines the various anatomical parts of the body by defining each organ system in relation to the body, and each sub-organ in relation to the organ-system containing it. The part-of hierarchy provides us with the taxonomic hierarchy for anatomical parts. A small section of the part-of hierarchy[8] and its graphical representation is shown in figure 7.

3.1.2 Material Flow Pathways

Material flow (e.g. the flow of glomerular filtrate) is represented by the *connected-to* relation. For example, the path of the filtrate in the kidney can be described as shown in figure 8. As can be seen from the figure, the material flow relation is specified at various levels of detail. The rationale for this multiple level description is provided later on in this section.

The anatomical knowledge that follows in the remainder of this section has been included to provide a fuller description of ABEL's knowledge base. However, this knowledge is currently not used by the program in its diagnostic reasoning.

---

8. The data are expressed in XLMS [Hawkinson80], which is briefly described in appendix 1.

Fig. 7. The part-of hierarchy

```
[body
   [urinary-system = (anat-entity*s "urinary-system")↑u
      [kidney = (anat-entity*s "kidney")↑u
         [cortex = (anat-entity*s "cortex")↑u]
          [medulla = (anat-entity*s "medulla")↑u]
          [nephron = (anat-entity*s "nephron")↑u
             [tubule = (anat-entity*s "tubule")↑u
                 [proximal-tubule
                    = (anat-entity*s "proximal-tubule")↑u]
               [loop-of-henle
                    = (anat-entity*s "loop-of-henle")↑u]
               [distal-tubule
                    = (anat-entity*s "distal-tubule")↑u]]
            [glomerulus = (anat-entity*s "glomerulus")↑u]]
          [collecting-duct
            = (anat-entity*s "collecting-duct")↑u]]
      [ureter = (anat-entity*s "ureter")↑u]
      [bladder = (anat-entity*s "bladder")↑u]
      [urethra = (anat-entity*s "urethra")↑u]]]
```

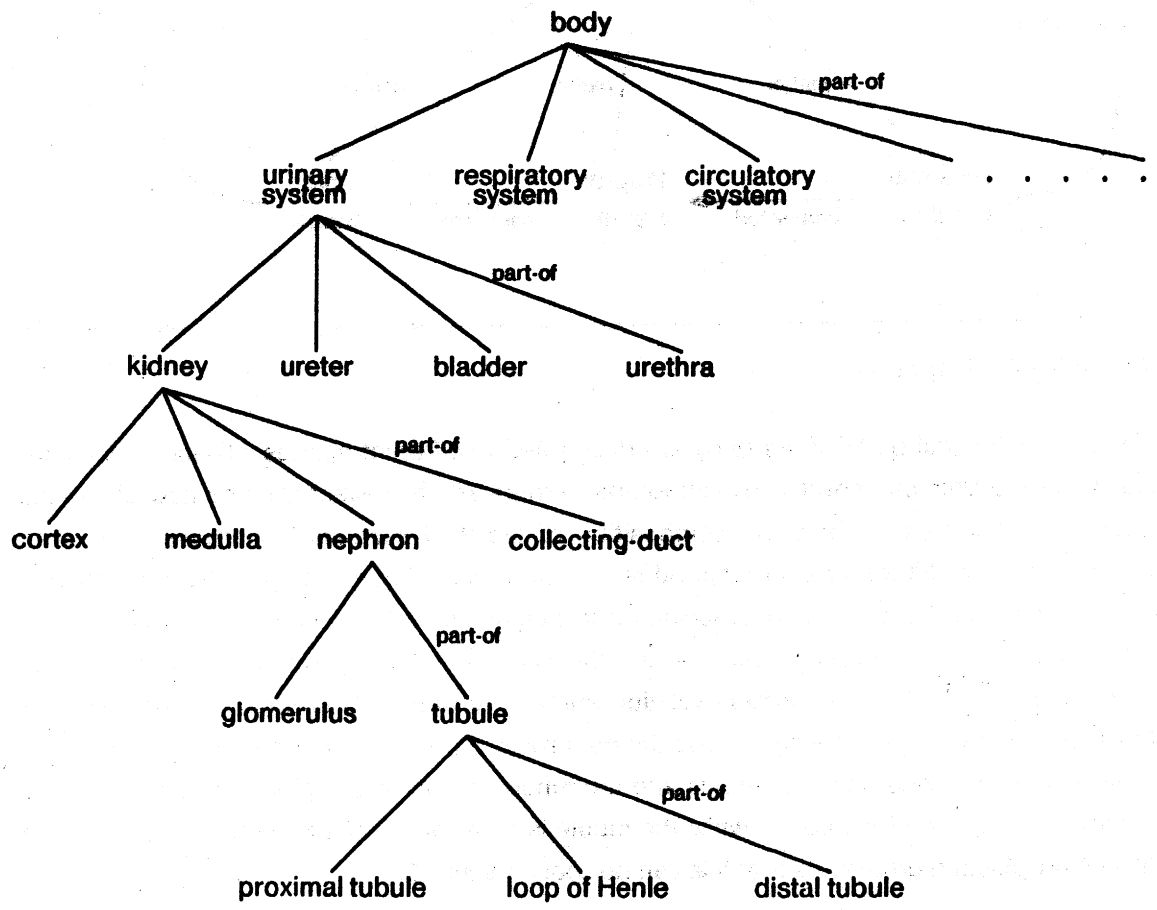
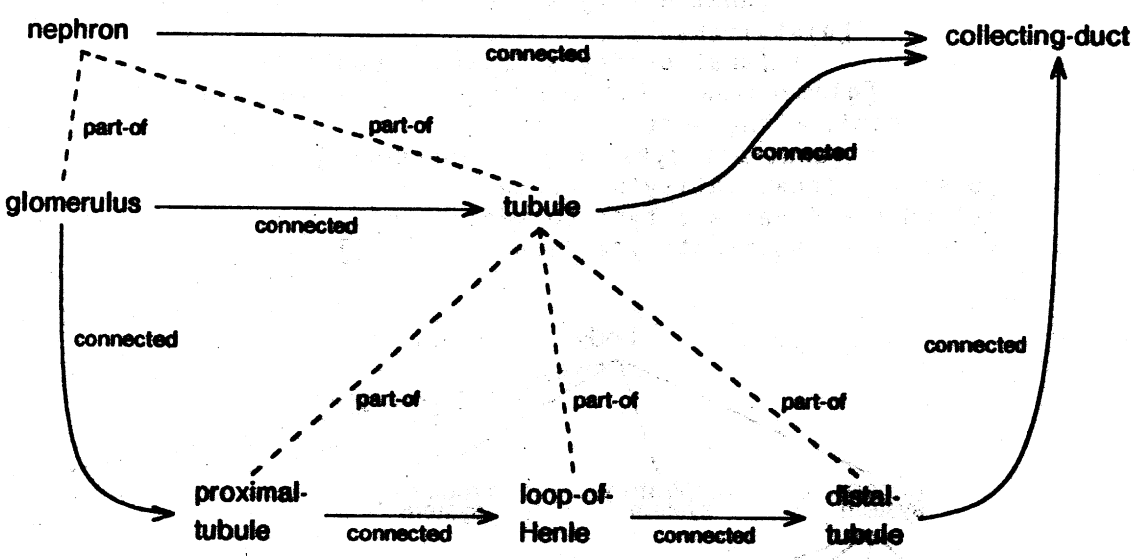


Fig. 8. Material flow relations

```
[((connected*b nephron)*e collecting-duct)]
[((connected*b glomerulus)*e tubule)]
[((connected*b tubule)*e collecting-duct)]

[((connected*b glomerulus)*e proximal-tubule)]
[((connected*b proximal-tubule)*e loop-of-henle)]
[((connected*b loop-of-henle)*e distal-tubule)]
[((connected*b distal-tubule)*e collecting-duct)]
```

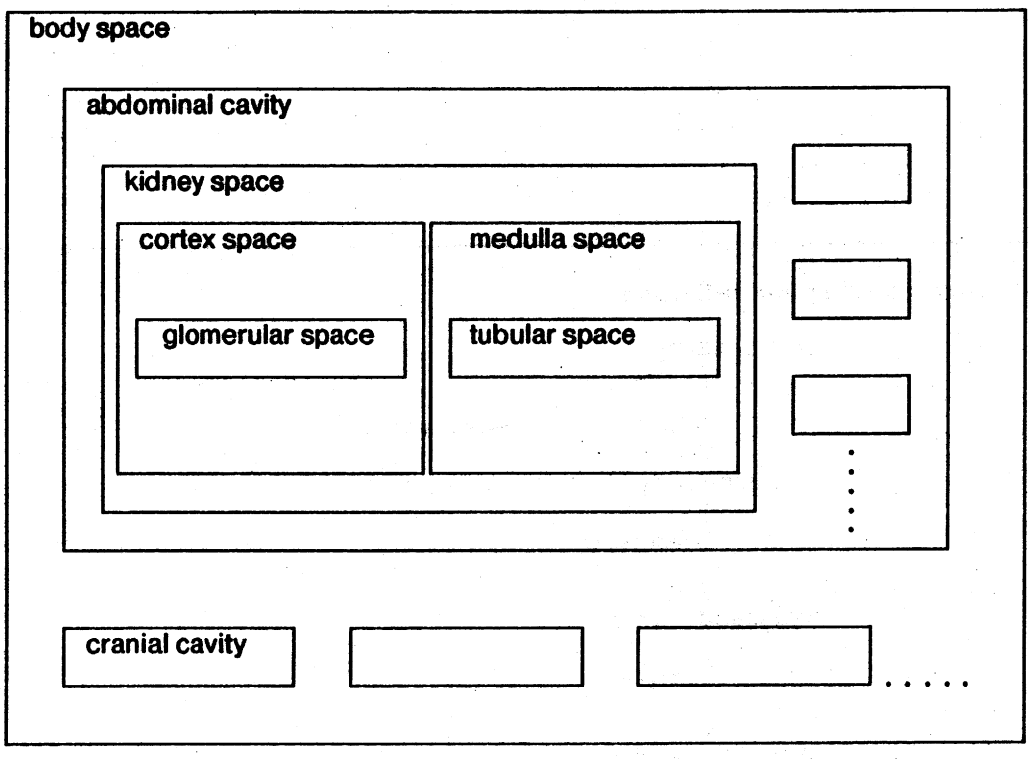


3.1.3 Anatomical Spaces

Various anatomical parts of the body are distributed in different spaces. These spaces are generally isolated from one another by membrane barriers which prevent the free flow of various electrolytes, proteins etc. Thus, the composition of the fluid surrounding organs in a given compartment can be different from that in other compartments. These general characteristics of the compartment can be useful in diagnosis and management of various diseases. Examples of such a compartment are the *cranial-cavity* and the *peritoneal-cavity*. Although the anatomical part-of relation and spatial containment relation are very similar, a distinction between the two must be made. For example, the cortex and the nephron are two different parts of the kidney and the nephron has two parts, the glomerulus and the tubule; however, the glomerulus is contained in the anatomical space of the cortex while the tubule is contained in the anatomical space of the medulla. A graphical representation of this can be seen in figure 9.

Fig. 9. The containment relation

```
[body-space
    [((contains*b body-space)*e cranial-cavity)]
    [((contains*b body-space)*e abdominal-cavity)]
    [((contains*b body-space)*e oropharynx-cavity)]
    [((contains*b body-space)*e thoracic-cavity)]
     .... ]

[abdominal-cavity
    [((contains*b abdominal-cavity)*e stomach-space)]
    [((contains*b abdominal-cavity)*e spleen-space)]
    [((contains*b abdominal-cavity)*e liver-space)]
    [((contains*b abdominal-cavity)*e kidney-space)]
     .... ]

[kidney-space
    [((contains*b kidney-space)*e cortex-space)]
    [((contains*b kidney-space)*e medulla-space)]]

[cortex-space
    [((contains*b cortex-space)*e glomerular-space)]]

[medulla-space
    [((contains*b medulla-space)*e tubular-space)]]
```

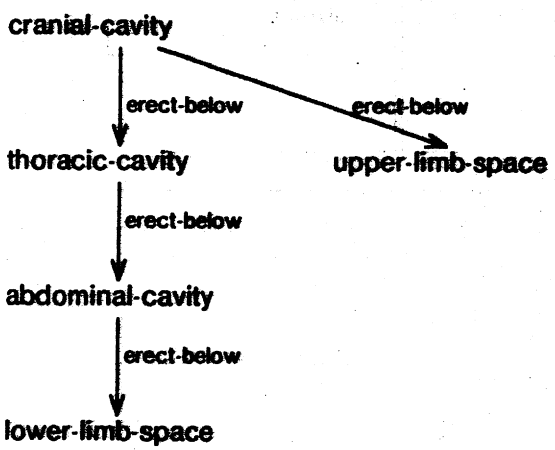


3.1.4 Miscellaneous Gross Anatomical Relations

A few additional anatomical relations are useful in common sense reasoning in medicine. An example of such a relation is the relative positioning of various anatomical spaces in supine position (lying face up in bed), erect position (standing up or ambulatory), etc. The use of this information can be illustrated by the following example. Let us consider a patient with nephrotic syndrome. A common symptom in nephrotic syndrome is periorbital edema (accumulation of fluid under the skin around the eye). In ambulatory patients, the periorbital edema can be observed only in the morning (after the patient has been lying down for some period of time); this accumulation of fluid can gravitationally move into other spaces once the patient has been up and around for some hours in the day. Thus the symptom is observable only in the morning and tends to disappear later in the day. Exactly an opposite effect can be observed in the case of pedal edema (accumulation of fluid in the feet) which tends to appear towards the evening and disappear in the mornings. This information can be used to explain away the absence of pedal edema in an edematous patient who is comatose. This information is encoded in the program with the use of positional relations as shown in figure 10.

We would like to note that the use of the anatomical knowledge in the current implementation of ABEL is limited to the use of anatomic taxonomy. However, we believe that the knowledge described here will be useful for further development of the diagnostic component as well as the therapy and prognosis components of the project.

Fig. 10. Gross anatomical relations

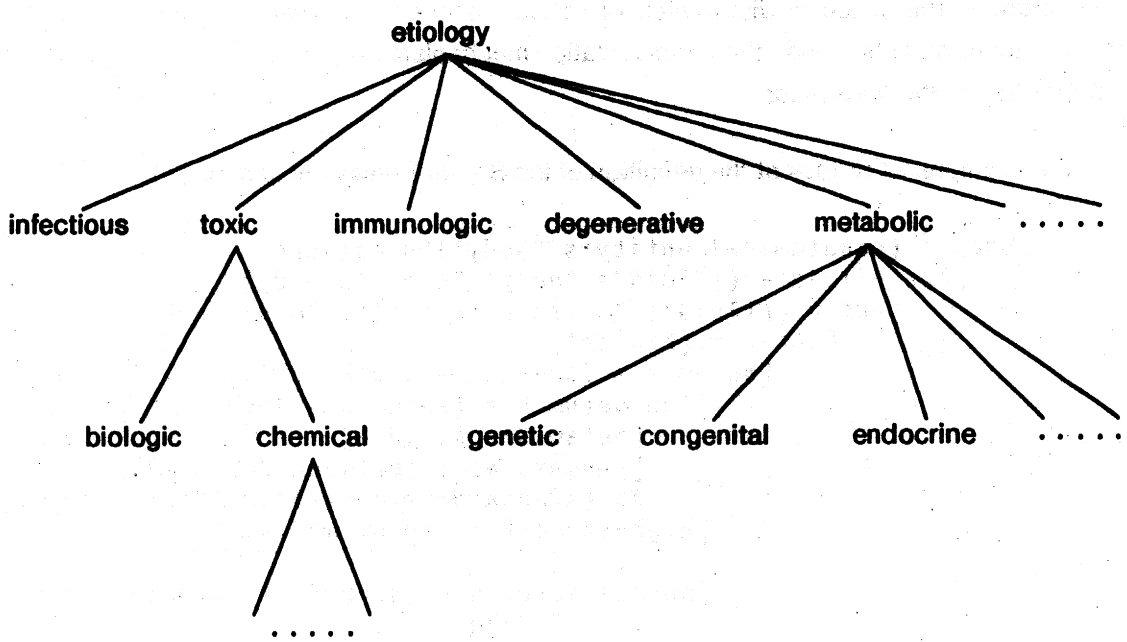


3.2 Etiological Knowledge

Disease categories are primarily organized around the organ systems; e.g., renal diseases, pulmonary diseases, liver diseases. In the previous section we have provided the basic framework of anatomical knowledge needed to provide such a categorization. The diseases of a given organ system tend to produce many symptoms associated with the loss of function of that system. For example, regardless of the cause of renal failure, all the diseases causing renal failure share common symptoms.

Fig. 11. Etiological hierarchy

```
[etiology = (medical-entity*s "etiology")
    [infectious = (etiology*s "infectious")]
    [immunologic = (etiology*s "immunologic")]
    [degenerative = (etiology*s "degenerative")]
    [toxic = (etiology*s "toxic")
        [biologic-toxins = (toxic*s "biologic-toxins")]
      [chemical-toxins = (toxic*s "chemical-toxins")]]
    [metabolic = (etiology*s "metabolic")
        [genetic = (metabolic*s ")]
      [congenital = (metabolic*s ")]
      [endocrine = (metabolic*s ")]]
    . . . . ]
```



Another important criterion for organizing diseases is the underlying mechanism causing the clinical disorder, i.e., the etiology of the disease. Similar to the anatomical categorization, the diseases with common etiology share symptoms common to the disease mechanism. For example, most infectious diseases cause fever. The taxonomy of etiologies in the program is shown in figure 11.

3.3 Physiological Knowledge

Knowledge about the normal functioning of the body and its adaptive response to abnormalities in body function plays an important role in the understanding and recognition of diseases. The need for this understanding is even more acutely felt in complex clinical settings involving the simultaneous presence of multiple abnormalities. Emphasizing this need, Dr. Jordan Cohen notes:

> *"The recognition of how common mixed disturbances are in complex clinical settings has served to emphasize the value of recognizing the limits of the physiologic response to simple disturbances, because frequently by reference to these limits one can recognize when a complex disturbance involving more than one simple abnormality is present."*
>
> *— New Concepts of Acid-Base Balance [Cohen77, page 1].*

In the physiological component of the program we have concentrated on the knowledge necessary in dealing with fluid, electrolyte and acid-base disorders. The physiological knowledge about fluids and electrolytes in the program deals with fluid compartments of the body and the distribution of body fluids in various fluid compartments, the composition of fluid in each compartment, the space of distribution of solutes, exchange of fluid and electrolytes between these compartments, and the homeostatic mechanism for regulating the quantity and composition of the body fluids.

For example, let us look at the definition of the Serum-Potassium concentration:

```
[body = ((anatomical-entity*s "body")*u patient)
    [body-fluid = ((fluid*s "body-fluid")*u body)
        [ecf = ((fluid*s "extracellular")*u body-fluid)
            [ecf-k = (K*u ecf)
                [serum-k = (concentration*u ecf-k)
                    [low-serum-k = (serum-k*f low)
                        [default*u #v !3.0]
                        [range*u #c !'(between 2.0 3.5)]
                        #s (standard-error*t !1.0,!1.0)]
                    [high-serum-k = (serum-k*f high)
                        .... ]
                    [normal-serum-k = (serum-k*f normal)
                        .... ]]]]]]
```

The above expression defines serum-K (serum potassium) to be the concentration of potassium

ion (K) in the extracellular fluid compartment (ecf), which is one of the components of the body fluid (body-fluid). The serum-K is further categorized as being either low (i.e., [(serum-k*f low)]), normal or high. Each of these categories is also associated with its default value, range and the acceptable amount of variance associated with its value (standard-error, in this case $\pm$1.0). The next example shows the encoding of the normal composition of the lower-GI-fluid. The lower-GI-fluid contains, in addition to water, Na, K, Cl and HCO_3. The quantities of these electrolytes and their variations are further specified in terms of the total quantity of the fluid. For example, the quantity of K is specified to be equal to 40.0 $\pm$10.0 meq/L of water in the lower-GI-fluid.

```
[lower-gi-fluid
    [water:u #c (quantity*u lower-gi-fluid)]
    [K:u #c (times*c water:u,!40.0)
         #s (standard-error*t
                (times*c water:u,!10.0),(times*c water:u,!10.0))]
    [Na:u #c (times*c water:u,!110.0)
         #s (standard-error*t
                (times*c water:u,!10.0),(times*c water:u,!10.0))]
    [Cl:u #c (times*c water:u,!80.0)
         #s (standard-error*t
                (times*c water:u,!20.0),(times*c water:u,!10.0))]
    [HCO3:u #c (times*c water:u,!40.0)
         #s (standard-error*t
                (times*c water:u,!10.0),(times*c water:u,!20.0))]]
```

In the previous three sections we have described the anatomical, physiological and etiological knowledge which, along with the temporal characterization, forms a basis for the taxonomic organization of diseases discussed in the next section.

3.4 Disease Knowledge

In studying the organization of medical knowledge about diseases Pople notes

> *"There are two conceptual frameworks that are used to organize medical knowledge, ... One of these employs the concept of causality or pathophysiology to establish a network of interrelated pathological states that might arise in the course of a disease. The other type of structure is the taxonomy of diseases, also called a "nosology", which is used to classify disease entities on the basis of anatomical locus, etiological agent, or other unifying principle."*
>
> — Structuring Medical Diagnosis [Pople81]

This section deals with the use of anatomical, physiological, etiological and temporal knowledge in defining a taxonomic disease hierarchy. With this taxonomic hierarchy in place, we will have completed the study of the basic medical concepts needed in ABEL for the description of disease pathophysiology.

A disease is defined in terms of its anatomical involvement, its temporal characteristic, its etiologic characterization and its pathophysiology. As each of the anatomic, etiologic, and physiologic knowledge is hierarchically organized, the locus of a disease along each of these dimensions can be selected at an appropriate level. A hierarchic organization for the disease definitions can then be derived from these loci.[9] For example, acute renal failure caused by nephrotoxic drugs could be specified as

```
[renal-disease = (disease*anat renal-system)
    [renal-failure = (renal-disease*object (urine-volume*f low))
        [acute-renal-failure = (renal-failure*tempch acute)
            [drug-induced-acute-renal-failure
              = (acute-renal-failure*etiology chemical-toxins)]]]]
```

The example above defines renal-disease to be a disease of the renal-system (anatomical locus). Renal-failure is then defined as a renal disease characterized by low urine output (physiological locus). Acute-renal-failure is defined to be renal-failure with an acute temporal characteristic, and finally, the drug-induced-acute-renal-failure is defined to be acute-renal-failure of chemically-toxic etiology. Note that each step of the above definition defines a disease which is further specialized by one of its primary characterizations. This provides a more specific placing of the diseases in the taxonomic hierarchy. In the next example we show how the disease definitions can be taxonomically organized along a single locus:

```
[GI-disturbance = (disturbance*anat gastrointestinal-system)
    [lower-GI-disturbance = (GI-disturbance*anat lower-gi-tract)]
        [disturbance-of-colon = (lower-GI-Disturbance*anat colon)]]]

[renal-disease = (disease*anat renal-system)
    [nephritis = (renal-disease*anat nephron)
        [glomerular-nephritis = (nephritis*anat glomerulus)]]]
```

As can be seen from the above two examples, the basic medical knowledge about anatomy, etiology etc., provides us with a framework for describing and organizing the disease hierarchy. We believe in the need for such a knowledge structure in the organization of any medical consulting program capable of expert level performance. However, we must note that this development is tentative and the details of the knowledge representation described above are likely to evolve considerably as its use in the diagnostic and therapeutic algorithms is better understood.[10]

---

9. Each hierarchy, such as the anatomic taxonomy, provides us with a tree structured partial order. The tree structure for the disease definitions is then derived from these partial orders.

10. In the current implementation of ABEL, this knowledge is used only for grouping different findings and diseases in diagnostic problem solving. However, the knowledge representation described in this chapter is capable of supporting a substantially wider variety of uses.

In the next section we will study the representation mechanisms for describing the causal (pathophysiological) knowledge relating different diseases.

3.5 Causal Link

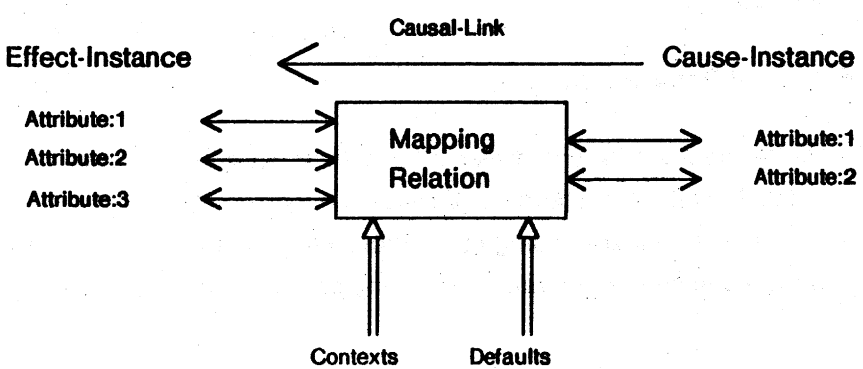
A causal link specifies the cause-effect relation between the cause (the antecedent) and the effect (the consequent) states. In the previous generation of programs (i.e., PIP, INTERNIST and GLAUCOMA), causal relations were described by links specifying the type of causality (e.g., *may-be-caused-by*, *complication-of*, etc.), and a number or a set of numbers representing in some form the likelihood (conditional probability), importance, etc., of observing the effect given the cause or vice versa. We believe that this simple representation of the relation between states is inadequate. The form of presentation of an effect and the likelihood of observing it depend upon various aspects of the presentation of the cause instance such as severity and duration, as well as on other factors in the context in which the causal phenomenon is manifested (such as the patient's age, sex and weight, and the current hypothesis about the patient's illness). To illustrate this, let us consider a (simplified) causal relation between diarrhea and dehydration. A rule-based description of this causal relation can be specified as follows:

```
IF diarrhea is severe,
    and its duration is greater than two days,
  THEN
    IF the patient has not received fluid replacement therapy
      THEN the patient is likely to have moderately severe dehydration
      ELSE the patient may have mild dehydration
```

From the above simple example, it is apparent that the conditional probability of observing dehydration and its severity and duration depend on the severity and duration of diarrhea and the fluid replacement therapy. Even this simplified example clearly demonstrates the need for information on how a cause relates to an effect, as well as other contextual information

---

Fig. 12. Schematic description of a causal link



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. A schematic description of a causal link and its representation in the data-base are shown in figure 12.

An example of the causal relation between total extracellular stores of potassium (ecf-K) and its serum concentration (serum-K) is described below.

```
[((caused-by*b ecf-k)*e serum-k)
    [context:1 #v total-ecf-water]
    [source:u
        [value↑u #c (times*c (value*c (value*u destination:u)),
                             (value*c context:1))]
      [start-time↑u #c (value*c (start-time*u destination:u))]]
    [destination:u
        [value↑u #c (quotient*c (value*c (value*u source:u)),
                                (value*c context:1))]
      [start-time↑u #c (value*c (start-time*u source:u))]]]
```

The causal relation between ecf-k stores and serum-k is specified by a causal link with cause (source) ecf-K and effect (destination) serum-k. The mapping relation describing this link is divided into two parts. The first part is associated with the source of the link and describes procedures for computing the attributes of the source (cause) given the attributes of the destination (effect), and the second part is associated with the destination given the attributes of the source. For example, the total quantity of potassium in the extracellular compartment (value of the source) is characterized as being the product of the quantity of the extracellular water (value of the context, total-ecf-water) and the concentration of the potassium ion in it (destination, serum-k).

Strictly speaking, it would not be appropriate to call all relations of this kind "causal," as some of the relationships are more matters of definition or association than cause. A more rigorous analysis, perhaps following the lines of [Rieger77], would further distinguish potential cause from actual, enabling conditions from true causations, etc. Such an expansion would, however, be orthogonal to our present argument, that any such link must connect several aspects of its source and destination.

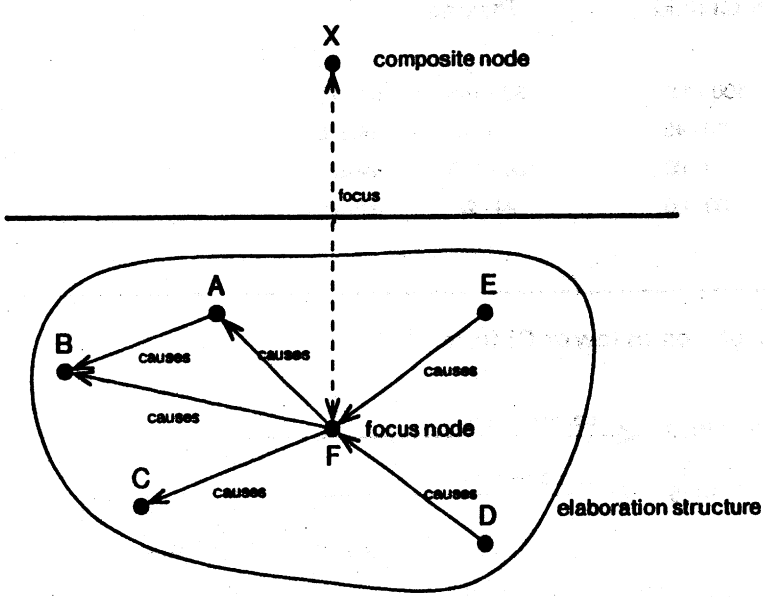
3.6 Multi-Level Causal Description

Medical knowledge about different diseases and their pathophysiology is understood to varying degrees of detail. Our understanding of medical expert reasoning also suggests that an expert physician may have an understanding of a difficult case in terms of several levels of detail. For example, *"serum creatinine concentration of 1.2 mg per cent"* is at a distinctly different level

than *"high serum creatinine"*,[11] and *"lower gastrointestinal loss"* than *"salmonellosis"*. For our program to reason at a sophisticated level of competence, it will need to share such a range of representations. In order to be effective the program must be able to describe the problem briefly yet still be able to take low level detail into consideration. We have attacked this problem by representing the program's medical and case-specific knowledge at five distinct levels of detail,[12] ranging from a pathophysiological level to a phenomenological level of knowledge.

The patient description developed here provides us with the ability to describe the patient's illness at various levels of detail. Each level of the 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 in the system 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, and other relevant attributes. A node is called *primitive* if it does not contain internal structure and is called *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*

Fig. 13. Schematic description of the node structure



11. A serum creatinine of 1.2 mg per cent can be interpreted in more than one way. For example we can assume this to be normal for a muscular male patient. But, for a average built female patient this could be an early indication of loss of as much as 1/3 of the renal function.

12. The number *five* does not have any medical or cognitive significance; it was chosen for purely engineering reasons.

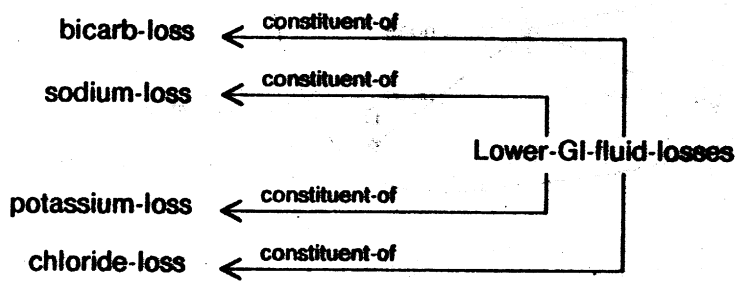
structure of the composite node. Figure 13 shows a schematic of the elaboration structure for a composite node labeled X. Nodes A through F and links between them form X's elaboration structure. Node X and F are connected together by a *focus link* making F the focus of the elaboration structure. The focus node identifies the essential part of the causal structure of the node above it. The collection of focal nodes acts to align the causal networks represented by different levels of the PSM. We note that very often a composite node and its focal description at the next level share the same name.[13] Nodes that do not play a role as the focal definition of any node at a higher level are called *non-aggregable nodes*. They represent a detailed aspect of the causal model which is subsumed under other nodes with different foci at less detailed levels of description.

To illustrate the description of a state at various levels of aggregation, let us consider the electrolyte and acid-base disturbances that occur with salmonellosis, which causes excessive loss of lower gastrointestinal fluid (lower GI fluid loss). In comparison with plasma, the lower GI fluid is rich in bicarbonate (HCO3) and potassium (K) and is deficient in sodium (Na) and chloride (Cl). The composition of lower gastrointestinal fluid and plasma are shown in figure 14. The loss

Fig. 14. Comparision of lower GI fluid and of plasma

	Lower GI fluid	Plasma	
Na	100 - 110	138 - 145	mEq/L
K	30 - 40	4 - 5	mEq/L
Cl	60 - 90	100 - 110	mEq/L
HCO3	30 - 60	24 - 28	mEq/L

Fig. 15. The loss of electrolytes in lower GI fluid



13. This is typical in English, where the level of detail of place names, for example, is often obtained from context and not encoded in the name used.

of lower GI fluid leads to the loss of corresponding quantities of its constituents as shown in figure 15. Therefore, an excessive loss of lower GI fluid without adequate replacement of fluid and electrolytes leads to a net reduction in the total quantity of fluid in the extracellular compartment (called *hypovolemia*). Because the concentration of K and HCO3 in the lower GI fluid is greater than in the plasma, there is also a corresponding reduction in the serum concentration of K (called *hypokalemia*) and HCO3 (called *hypobicarbonatemia*) in the extracellular fluid. Finally, because the concentration of Cl and Na in the lower GI fluid is lower than that in the plasma, there is corresponding increase in the concentration of Cl (called *hyperchloremia*) and Na (called *hypernatremia*) in the extracellular fluid. A graphic representation of this information at the next higher level of aggregation is shown in figure 16. Figure 17 shows the aggregation of this information along with some additional causes and consequences of lower GI loss at the next more aggregate level of detail. Hypobicarbonatemia is interpreted as metabolic acidosis at the next higher level of detail. Note that the hypernatremia and hyperchloremia have not been encoded at this level.[14] The hyperchloremia was not encoded because it is not clinically significant. The hypernatremia, however, is not encoded because it is not a common finding in the presentation of lower GI loss. The lower GI loss at this level is a non-aggregable state and therefore does not have a focal aggregation at the next level above. Figure 18 shows the

Fig. 16. Consequences of lower Gi loss described at next higher level

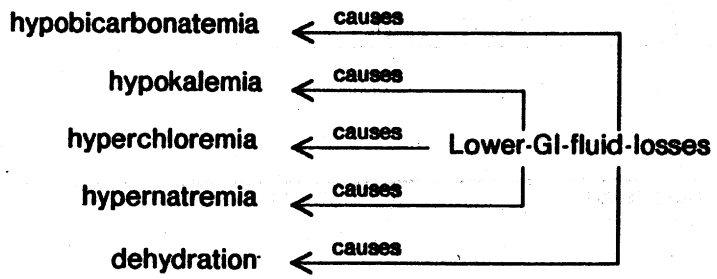
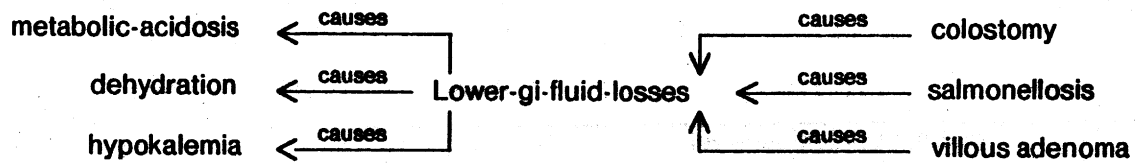
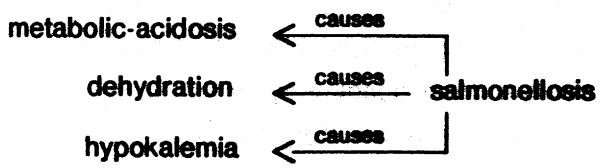


Fig. 17. Lower Gi loss expressed at an intermediate level



14. The causal knowledge described here is encoded by hand, and represents the program's general medical knowledge. A similar multi-level description built by the program to describe a specific patient illness will be discussed in chapter 4.

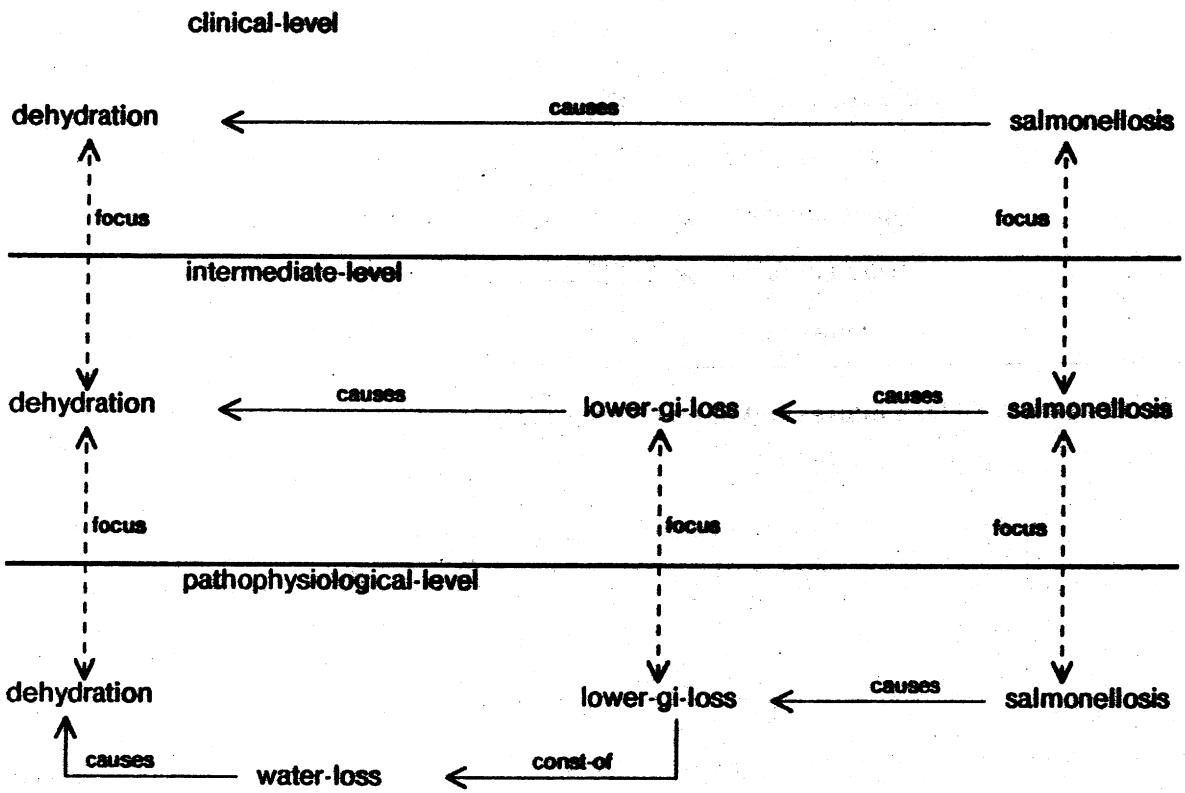
Fig. 18. Salmonellosis and its consequences expressed at the clinical level



description of the aggregate effects of salmonellosis (one of the causes of lower GI loss).

Links can be categorized into two types, as nodes are: the *primitive* links and the *composite* links. To illustrate the concept of elaborating causal links to form a causal chain, let us consider the causal relation between salmonellosis and dehydration shown in figure 19. The causal mechanism of dehydration caused by salmonellosis can be elaborated as follows: salmonellosis causes lower GI loss, which in turn causes dehydration. Expressed at the next level of greater detail, the lower GI loss leads to water loss which results in reduction in the extracellular volume.

Fig. 19. Layered description of link: salmonellosis causes dehydration



The state of reduced extracellular volume is called dehydration.

Because the causal relations specified by links are not guaranteed to be true under all circumstances (they represent strong associations, not logical truth), the validity of deductions degrades with every additional intermediate link. That is, a causal pathway containing a large number of links is less likely to be valid than one using only a few links. Therefore, in order to explore a large diagnostic space, we must reduce the lengths of commonly occurring chains of causal relations. One way of achieving this is through the multi-level description proposed in this chapter. The multi-level description scheme allows us to aggregate the diagnostic space to a level where each link represents an aggregate causal phenomenon covering large distances and thus minimizing the possibility of error in the deduction.

However, the multi-level description proposed above can not solve this problem completely. For example, there are situations where all the intermediate nodes in a given causal chain cannot be suppressed due to limited number of levels of description. Stated differently, because of the fixed number of levels in the multi-level description, the programs ability to aggregate causal