PROCEEDINGS

AAMSI CONGRESS 1986

EDITED BY
ALLAN H. LEVY, M. D.
BEN T. WILLIAMS, M. D.

AAMSI CONGRESS 86

Sponsors:
American Association for Medical Systems and Informatics
American Association for Artificial Intelligence
American College of Medical Informatics
California Medical Association
Federation of Computer Users in the Medical Sciences
George Washington University School of Medicine and Health Sciences
Illinois Conference on Medical Information Systems
International Health Evaluation Association
Kaiser-Permanente Medical Care Program
MUMPS Users' Group
National Academy of Practice in Medicine
National Institutes of Health
Division of Computer Research and Technology
Office of Biometry and Field Studies, NINCDS
National Library of Medicine Lister Hill Center for Biomedical Communications
Opus Unitas
Symposium on Computer Applications in Medical Care
Veterans Administration Medical Information Systems
University of California-San Francisco School of Medicine, EPME
University of Missouri-Columbia Medical Information Sciences Group
University of Southern California Medical School, CME

Supporting Sponsor:
American Association for Medical Systems and Informatics

Editors:

ALLAN H. LEVY, M.D.
Departments of Medical Information Science and Computer Science
University of Illinois, Urbana

BEN T. WILLIAMS, M.D.
Departments of Medical Information Science and Pathology
University of Illinois, Urbana and Mercy Hospital, Urbana

Proceedings of the AAMSI Congress 86, Anaheim, California, May 8-10, 1986

Volume 4
ISSN 0737-4194

American Association for Medical Systems and Informatics, Publishers
Suite 700, 1101 Connecticut Avenue, N.W., Washington, DC 20036 202/857-1189

Printed in the United States of America
Dependable Printing Company, Inc.

Preface

On the Use of the Future Imperfect Tense

Medical informatics is both a science and a hope. As a science, it seeks to develop an orderly body of knowledge concerning the origin, acquisition, classification, analysis, and dissemination of information dealing with biology, health and disease. As a hope, it seeks to use knowledge to improve health care delivery, the treatment of illness, and the quality of biomedical research.

Some papers in this volume report on advances in the science of medical informatics and others on the uses of information technology as a practical tool for education, medical practice and biological research. Both types of papers make important contributions to the store of knowledge and to the advancement of society. Often, there is no sharp line between the two types of reports. A careful report of an implementation of an information system will have all the elements of a sound research report: it will state the aim, how the plan was carried out, what data were collected in order to observe the effects, and how these data were evaluated. Only then, will the report draw conclusions about the possible implications of the the work. Many of the papers in this volume are consistent, at least in part, with this simple style.

However, this is not always the case. In preparing this volume, we encountered a few examples--fortunately not represented in the pages to follow--of what we shall irreverently term the excessive use of the future imperfect tense. Here, the worker failed to distinguish between what the results actually were and what the results were intended to be. In fact, the failure was usually not so much an error in analysis or a misinterpretation of the observations, but actually a failure even to record the critical observations that would have permitted an interpretation.

Whether one is studying enzyme kinetics, devising a better patient record, or doing basic research in computer science, the fundamental principles are the same: hypothesize, perturb, observe, analyze and repeat. It is especially important to keep the simplicity of the scientific method in mind. Our tools are more sophisticated, and our choices more abundant, but essentially, methods of study still embrace the same straightforward principles. Now that informatics has become such a pervasive part of biomedicine, it is critically important that this be kept in mind, in order that we not contribute to information pollution, the prevention of which should be one of our principal aims.

We acknowledge with great appreciation the assistance of Mrs. Mary Ann Steiner. Her attention to every detail made this Congress and these Proceeedings a reality. Our colleagues on the Program Committee always responded, and the staff of the AAMSI office did their jobs efficiently and well.

Allan H. Levy, M.D.
Ben T. Williams, M.D.
Urbana, Illinois

CONTENTS

Proceedings

of the congress on

Computer Applications in Medicine and Health Care

COORDINATING CLINICAL & PATHOPHYSIOLOGIC KNOWLEDGE FOR MEDICAL DIAGNOSIS

Ramesh S. Patil

Laboratory for Computer Science
Massachusetts Institute of Technology
Cambridge, MA 02139

Complex clinical problems often result from the interplay of multiple diseases where the presentation of one disease may be altered or masked by a concomitant disorder. Most existing programs for medical diagnosis however, assume that only one disorder is present. They cannot, therefore, deal adequately with the presence of multiple disorders. In this paper I present an experimental program called ABEL designed to address this problem. ABEL uses its detailed knowledge of disease pathophysiology to deal correctly with multiple disorders. It uses clinical knowledge to guide its pathophysiologic reasoning for efficient diagnosis. Finally, it uses quantitative relationships between causes and effects to sort out interactions between diseases and identify partially masked disorders.

Introduction

Last two decades have seen a rapid growth in the field of *Artificial Intelligence in Medicine* (AIM). The first decade of research in this field culminated in many promising programs, among which are INTERNIST-I [1], the Present Illness Program [2], CASNET/Glaucoma [3], and MYCIN [4]. These programs represent the first efforts in the use of sophisticated AI techniques in medical diagnosis, and therefore have come to be characterized as *first generation AIM programs* [5]. Each of these programs made contributions in one of many dimensions of the art of medical diagnosis by computers. For example, INTERNIST-I and PIP successfully demonstrated an ability to quickly focus on a small set of diagnostic hypotheses in broad domain such as internal medicine. CASNET made effective use of pathophysiologic knowledge and MYCIN pioneered in the areas of dealing with uncertainty, explanation generation, and tutoring [6]. Most of these programs have, in some trial, been judged to match expert physicians in their competence [7] and thus have clearly established the feasibility and usefulness of computers in medical diagnosis.

As experience was gained with these programs it became clear, however, that these programs could not adequately exploit the rich underlying causal structure of disease knowledge, their temporal character or severity of disease presentation, and had a limited capability to cope with the wide variations in clinical picture that may occur in a range of patients suffering from the same disease. They may, therefore, fail to uncover the subtle discrepancies which can help to distinguish between alternative diagnoses. Further, they have no ability to recognize how one disease alters the clinical presentation of another or to take into account the effect of therapeutic interventions in a patient for related or unrelated problems. Given the frequency with which such situations arise, this represents a serious shortcoming. This lead to a new set of research efforts towards the development of a *second generation* of computer programs such as ABEL [8], CADUCEUS [9], NEOMYCIN [10], and MDX [11] aimed at synthesizing successful features from the first generation programs and augmenting them with new capabilities for remedying their weakness. In this paper, I will be describing an experimental program (ABEL) designed specifically to address the problem of multiple interacting diseases. Although the program described here is devoted to the narrow domain of acid-base and electrolyte disorders, the issues we have dealt with are generic in character.

The First Generation Approach to Multiple Disorders

Nearly all early programs that attempted to deal with multiple disorders limited themselves to diseases without overlapping findings. These programs assumed that all hypotheses are competitors and focused

their efforts on identifying the single most likely diagnosis [1, 2]. Only after the first diagnosis was confirmed did they attempt to make a second diagnosis based on the residual findings. This process was repeated for so long as there were any findings not accounted for by an already-confirmed hypothesis. Such a sequential approach suffers from two serious deficiencies. First, the program has no way of recognising that more than one problem exists until after it has already succeeded in identifying the first problem. Because of this inability to take into account the presence of additional disorders the program is initially forced to attribute all observed findings to the first diagnosis it is trying to establish. As a result, findings that are not in fact relevant to the primary diagnosis can easily confound the diagnostic process. Thus, in a patient suffering from chronic glomerular nephritis and acute myocardial infarct, the program will mistakenly try to attribute the manifestation of both diseases to the leading contender. As a result, for example, the program will find the hypothesis of chronic glomerular nephritis far less compelling because of its inability to account for severe chest pain.

A heuristic solution (called partitioning heuristic) was developed and used in the INTERNIST-I program. This heuristic is based on the assumption that co-existing disorders should, in general, account for a larger set of observed findings than either one alone. This heuristic was used to separate the active diagnostic hypotheses into two groups. First, the *competing group* containing hypotheses which competed with the leading hypothesis, i.e., explained only a subset of findings explained by the leading hypothesis. Second, the *complementary group* containing hypotheses that complemented the leading hypothesis, i.e., explained some finding(s) not explained by the leading contender. The program then focused its diagnostic activities to the competing group setting aside the hypotheses in complementary group for later consideration. The partitioning heuristic, however, fails to deal correctly with situations in which a patient has two diseases whose findings overlap appreciably or when the two diseases interact.

The second problem arises because, in the presence of multiple disorders, it is not clear when a program can reasonably conclude that some finding has been successfully accounted for. This question is crucial because in the absence of quantitative information on severity a program must allow a finding to be either completely accounted for by a disease or not at all. Furthermore, a finding that has already been accounted for by one diagnosis can either be allowed to continue to lend support to additional diagnoses or it may not. Either of these choices lead to problematic behavior. The first choice leads the program to continue its diagnostic activity interminable in pursuit of ever-more implausible combinations of diagnoses that would account in new ways for findings that have already been accounted for adequately. The second choice, on the other hand, can often prevent the program from correctly diagnosing a co-occurring disease which shares a significant fraction of its findings with an already confirmed diagnosis.

Reasoning with Multiple Interacting Diseases

To illustrate the rich character of clinical reasoning involved in the diagnosis of multiple disorders, let us consider the case of a patient suffering from diarrhea and vomiting who is hypovolemic, hypokalemic and has serum pH within the normal range. Diarrhea and vomiting both cause substantial loss of body potassium. Thus taken together, they compound the effect of each other on observed hypokalemia. Conversely, diarrhea results in loss of alkalis and vomiting results in loss of body acids. Therefore, taken together they tend to offset the effects of one another on serum acidity. For the sake of example, let us also suppose that we know about the vomiting but are not aware of the diarrhea. In such a situation, the observed hypokalemia is too severe to be properly accounted for by the vomiting alone, it cannot be considered to be a complete explanation for the observed severity of hypokalemia. Given this fact, a program must consider vomiting either as not being responsible for hypokalemia or only partially responsible for it. If vomiting is not held responsible, the further reasoning, although erroneous, is quite simple – we must find the actual cause for hypokalemia. If vomiting is partly responsible, however, we must be able to determine the part of hypokalemia that can be attributed to vomiting and identify the part that still remains to be accounted for. Furthermore, when a second cause for hypokalemia is identified, we must be able to judge how well the two causes taken together explain the observed hypokalemia.

Let us pursue the above example further. Given that the patient is suffering from vomiting severe enough to cause significant hypokalemia, one should anticipate that the patient is also suffering from alkalemia

(loss of potassium and gastric acids share a common pathway – the vomitus). However, the patient is not alkalotic. The programs discussed above will treat this situation erroneously; they will use the absence of anticipated alkalemia as evidence against vomiting and lower their belief in it. They will, therefore, fail to identify the second disorder, namely diarrhea, which is surreptitiously masking the effects of vomiting on serum alkalinity. Furthermore, they would consider serum acidity within normal range to be "normal" and, thus, not requiring any explanation. A program that allows a proper accounting for the findings will, however, attribute only a part of hypokalemia and hypovolemia to the vomiting and will be able to identify an as yet unknown factor compensating the effects of vomiting on serum acidity. It will thus be able to hypothesize the presence of a second disorder which in the absence of vomiting should lead to hypokalemia, hypovolemia and acidosis.

To capture this richness of medical knowledge and clinical reasoning, a new approach must be taken. A program must allow explicit representation of multiple interacting disorders within a single composite hypothesis. Such a composite hypothesis must integrate all features of the individual diseases and their interactions and attempt to explain the observed manifestations of a patient's illness in entirety. It must incorporate an explicit notion of severity and quantity, and use pathophysiological knowledge to identify and sort out potential interaction among co-occurring diseases. To reason efficiently, a program must represent its knowledge at multiple levels of detail that allows it to reason about the same problem at a shallow associational level that captures the clinical intuitions and at deeper causal level that can model in detail the pathophysiology of disease processes.

Organisation of Medical Knowledge

The basic medical knowledge in ABEL consists of hierarchical representations of anatomical, physiological, etiological and temporal knowledge. A disease is then characterised in terms of its anatomical involvement, its temporal character, its etiologic origin and the functional derangement resulting from it. As each of the anatomic, etiologic, and functional knowledge is organised in a taxonomic hierarchy, the projection of a disease description along each of these dimensions can be used to uniquely derive a lattice structure (based on the subsumption relation [12]) so that a general description of a disease or clinical state appears above more specific descriptions. The disease descriptions are then augmented using causal relationships.

The causal knowledge in the program, as indicated earlier, is organised at several levels of detail. At the shallowest level this knowledge is in terms of diseases and their clinically observable manifestations. At the deepest level this knowledge includes detailed biochemical and pathophysiologic mechanisms which provide quantitative relations among normal and abnormal physiologic parameters and processes. Additional information is also provided to describe the connection of knowledge at one level to that at adjacent levels.

The causal knowledge at each level of detail is organized in terms of nodes and links. Nodes are clusters of information that describe physiologic and clinical states. Nodes are linked to one another by causal links or by links which describe associations when underlying causal mechanism is not clear. Causal links may connect a node describing a disease or a clinical state to one or more nodes which describe their effects. They specify the relationship among the severity, duration, and other relevant aspects of the cause and the effect nodes, that is, given a cause and an effect node it is possible to compare the two for causal consistency. Furthermore, reasoning may also be carried out in the forward or the reverse direction; a cause may be used to predict the effects or an effect used to deduce the necessary severity and duration of a cause. Additional information is also provided to permit the combining of separate effects into a joint one when multiple causes are present or suspected.

Multi-level representation of nodes allows the knowledge base to describe a high level node (called a *composite node*) in terms of a network of states and causal relations at the next lower level (Figure 1). One of the nodes in this causal network is designated as the *focus node*. The focus node identifies the essential part of the causal structure (called the *elaboration*) of the node above it. Indeed, the collection of focal nodes act to align the causal network representing the medical knowledge at different levels of

detail. Nodes that do not play a role as a focal definition of any node at a higher level are called *non-aggregable* nodes. They represent the detailed aspects of causal model introduced at the given level which was subsumed under other nodes with different foci at less detailed levels of description. Finally, nodes which are not described at the next level of detail are called *primitive nodes*. Such a situation arises when either the pathophysiology of a given state is not available, or it is not medically relevant.

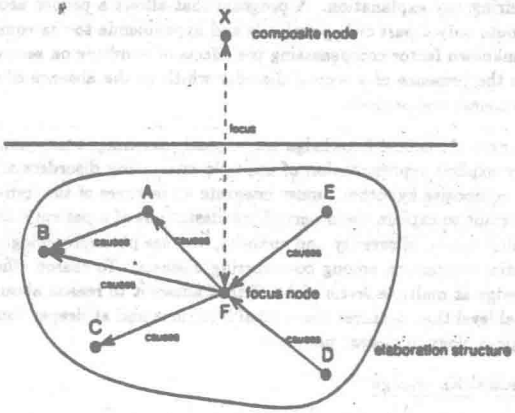

Figure 1. Schematic Description of a Composite Node Structure.

Multi-level representation of links allows the knowledge base to describe a high-level relation between two clinical or pathophysiologic states at next more detailed level using a chain of causal relations. Similar to nodes, links described in such a manner are called *composite links*, and links which do not contain such structure are called *primitive links*. A schematic causal relation described at multiple levels of detail is shown in Figure 2.

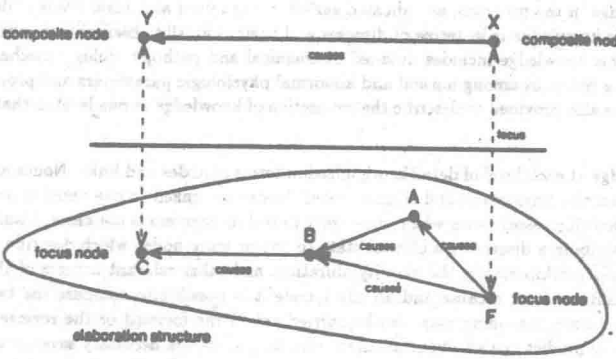

Figure 2. Schematic Description of a Composite Link Structure.

One important function of diagnostic reasoning is to causally relate the disease and symptoms observed in a patient. The causal pathways associated with links play a key role in elaborating clinical level description to detailed pathophysiologic level, whereas the causal network associated with a node play a central role in identifying clusters that can be meaningfully aggregated in developing a coherent diagnosis.