

EXPERT SYSTEMS IN MEDICAL DIAGNOSIS: A DESIGN STUDY IN
DERMATOPHYTE DISEASES

By

OH, KYUNG NA

B.S., San Jose California State University, 1975

A MASTER'S REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Computer Science

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1983

Approved:


Major Professor

LD
2668
R4
1983
.04
c.2

A11202 244431

ii

ACKNOWLEDGMENTS

I wish to express my sincere thanks to my advisor, Dr. Roger T. Hartley, for his valuable guidance and suggestions to make this report possible to complete. Thanks are also given to my patient family and to the other members of my committee, Dr. Rodney M. Bates and Dr. Paul S. Fisher.

TABLE OF CONTENTS

	Page
LIST OF DIAGRAMS	iv
CHAPTER	
1. INTRODUCTION	1
2. THE PURPOSES OF COMPUTER-AIDED MEDICAL CONSULTATION	4
3. EARLIER FORMALISM METHODS	7
3.1 Statistical and Logical Approach	7
3.2 Mathematical Approach	8
3.3 Flowchart Algorithm	9
3.4 Decision Theory	10
3.5 Pattern Recognition	11
3.6 Summary	11
4. ARTIFICIAL INTELLIGENCE (AI) METHOD AND IT'S GOALS	13
5. HYPOTHESES IN MEDICAL REASONING	16
6. A CASE STUDY (DERMAS)	19
6.1 DERMAS as a Task Environment for AI	19
6.2 Domain of DERMAS	20
6.3 Major Components of DERMAS	22
6.4 Knowledge Base of DERMAS	25
6.5 Logic of DERMAS	33
6.6 Consultation Process of DERMAS	35
6.7 Evaluation and Future Work of DERMAS	50
7. PERFORMANCE AND ACCEPTABILITY OF AI CONSULTATION SYSTEMS ...	52
8. FUTURE WORK	54
BIBLIOGRAPHY	55
APPENDIXES	58
A. Glossary of Medical and Taxonomic Terms	A-1
B. Anthrophilic Dermatophytes	B-1

LIST OF DIAGRAMS

	Page
DIAGRAM	
1. The Structure of DERMAS	26
2. A Portion of Knowledge Base of DERMAS	31
3. Flowchart of Consultation Process	36
4. Some Examples of Dialog by QA Program	38
5. An Example Generated by the Explanation Program	41
6. A Typical Frame Based on a Disease Entity	44
7. Generated Diagnosis and Treatment Recommendations	47

CHAPTER 1

INTRODUCTION

During the past twenty years, computer scientists and physicians have been interested in the use of computers as an aid to both clinical decision making and management. Since then, a variety of programs and papers have been written about the potential of computers to help in the diagnoses of diseases, planning of treatments, calculating the correct dosages of medication, monitoring intensive care patients and management of patients' records and histories in an automated fashion. Most of these projects are restricted to a specific subfield of medicine in practical usage and are still at an experimental stage, but they offer promise in the near future.

There is a variety of techniques that have been applied to these programs; however, they all share a single characteristic - they are highly structured, adopting some kinds of formalism to capture clinical expertise in the computer. While physicians often consider that their clinical reasoning process is more "art" (3) than science, and that a computer may never be able to capture the "art" of a skilled physician, it can be programmed to deal with the scientific and analytic aspects of a skilled physician. Therefore, the purpose of this paper is to attempt to mimic as closely as possible with a computer, the reasoning process used by a skilled physician in making a medical diagnosis.

Even though there have been attempts to simulate the physician's reasoning process, until not too long ago, remarkably little was known about the medical reasoning process as a cognitive science (3). This lack of knowledge was due to a lack of suitable tools to study the

physician's reasoning process. It has been clear that previous formalisms before Artificial Intelligence (AI) methods were introduced, were not suitable to handle this kind of investigation. Because there was no single, formal approach which could accommodate the knowledge, experiences, and common sense of a skilled physician in a real world situation, AI people have been studying the medical diagnosis process as human problem solving (3, 12, 18). The study of human problem solving is the study of natural intelligence; in fact, this is one of the central activities of Artificial Intelligence.

In order to accomplish a computer-based capability in this problem solving domain comparable to that of a skilled physician, development of a sophisticated model of synthetic reasoning based on AI methods is necessary to emulate those aspects of the physician's behavior. Medical reasoning is a complex problem requiring large knowledge bases that contain complex interrelationships including time and functional dependencies. The knowledge of medical domain is incomplete, thus this must be extensible. Careful study and thorough understanding of its task, high level of representation and utilization of knowledge base will heighten its performance. It is particularly important to find effective ways of matching the knowledge base and the reasoning strategies to the complexity of the task domain. Therefore, it is inevitable that more computer medical consultation systems are handled by collaboration of AI and medical people.

Even though there are a number of people who are experts as both physicians and computer scientists, the number of such people is very small. Also, since it takes such a long period to develop such

specialties in both fields, the number will remain small. Thus, collaboration is essential to create a computer consultation system.

In this paper, the following subjects will be examined concerning computer and a medical consultation:

- 1) The purpose of computer-aid in medical decision making,
- 2) Non-AI formalism methods used in medical diagnosis,
- 3) AI method and its goals,
- 4) Hypotheses in medical reasoning,
- 5) A design study, and
- 6) Evaluation of existing AI system's performance and acceptability.

The design study is about a diagnostic system dealing with dermatophyte fungal infections. The particular subfield of medicine has been chosen as the domain of the design part of this report project because there is no existing AI consultation system dealing with the same problem domain. In the design part, the author attempts to design a diagnostic system which can emulate a physician's reasoning process utilizing AI methods; including development of the knowledge base, the consultation mechanisms, the control strategies, and the consultation process.

CHAPTER 2

PURPOSES OF COMPUTER-AIDED MEDICAL CONSULTATION

Because of the rapid growth of biomedical information, the available body of knowledge is far greater than any individual can possibly assimilate; thus, it is necessary to find a solution to cope with this problem. For instance, a study (12) shows that there are approximately one million facts as the core body of information in general internal medicine and two million facts in the case of information in the subspecialties of internal medicine. Even if the study somewhat over-estimates the number of available information, it is impossible to use all the possible information and the knowledge in the actual process of making a diagnosis. Therefore, physicians have a tendency to specialize and become more dependent upon other specialists when the presented case is a complex problem requiring knowledge outside of their own specialty or specialties.

Correct diagnosis depends on the physician's ability to synthesize the available information and apply it appropriately to the problem solving task.

Keeping the earlier mentioned figure in mind, several conclusions can be drawn about the reasons for the need for computer assistance in the medical domain. They are:

- 1) Accuracy of diagnosis will be improved, because medical data will be available in a complete, organized, and systematic fashion; and the danger of human memory lapse will be avoided. Data stored in the computer will be standardized and confusion of a variety of terms can easily be eliminated through a data dictionary.

- 2) Consistency of performance will be improved since the machine's performance is not influenced by lack of time, fatigue, emotions, or sicknesses. Also, there will be less tendency to confuse similar cases; therefore, overall reliability will be improved.
- 3) Since the computer can calculate the value of inconveniences, costs, risk of possible treatment, and the required length of therapy (18), patients can benefit from this information before making definite plans for the next step of therapy.
- 4) Well written computer programs help to improve our understanding of the structure of medical knowledge and the clinical reasoning processes (18). This understanding can be used as a powerful tool in medical teaching to guide medical students in decision making through a structured method rather than an unstructured method based on observation and imitation of other physicians' reasoning processes.
- 5) The immediate accessibility of a computer improves efficiency. Since most physicians have a very tight schedule, a physician is not always able to consult with needed specialists.
- 6) Computers, with their capability of handling massive quantities of information, can reduce the burden on the physician's memory which is now imposed by the huge number of important medical facts. It will be an extension of resources of the physician's information and knowledge (10). This can be particularly advantageous in gathering information from patients without overlooking any small facts which may be significant.

- 7) Overall health care quality can be improved as computers become more readily available at a reasonable cost as part of a medical consultation system.

CHAPTER 3

EARLIER FORMALISM METHODS

As was discussed briefly in the introduction, several formal methods have been applied in the medical decision making process. They are used mainly in diagnostic reasoning although some methods have been intensively applied to therapy problems. In this chapter, various formal methods will be reviewed to clarify the differences between non AI methods and AI methods for better understanding of AI methods.

3.1 Statistical and Logical Approach

As early as the mid 1940's, statistical hypothesis testing methods were used, mostly for screening in radiology (18). It was done simply by calculator computation. Then, logical schemes (10, 18) followed for matching a patient's symptoms to prior stored diagnoses, using cards for sorting and matching. One of these methods used a boolean combination of descriptors to match the presented symptoms to previously stored information.

By the first half of the 1960's, a combination of statistical and logical techniques were introduced utilizing computers (3, 10, 18). The best known of this type is based on Bayesian approach (8, 18, 19). Even some AI methods adopted this approach for a rating methodology in various ways. For example, a medical consultation program designed to assist physicians in diagnoses of acute renal failure (8) uses Bayesian statistical approach extensively.

Initially, complete information about signs, symptoms and laboratory data are collected, and then the information is processed by

utilizing Bayesian techniques. The probabilities of candidate diseases change whenever new data are processed. The final diagnosis is reached when all of the predetermined set of questions is asked. After each question, the reply is in the form of a ranked list of probabilities of the possible diseases from highest to lowest. Also the risks and the benefits of different therapies, and the result of certain therapies are estimated based on probabilities. The effects of therapy are divided into three categories of improved, no change, and worse based upon probabilities.

The major attraction of this approach is precision, being able to compute the exact probability of a specific disease or a symptom based on observations and data regarding its frequency. It is a mathematically sound approach and has shown a high level of correctness in some domains in which it has been tested, such as cases of testing gastrointestinal diseases (18).

Various statistical and logical methods, or a combination of both, depended heavily upon large data bases of reliably diagnosed case histories for diagnostic decision making (10, 18). When the set of information about patients was well standardized (or clearly specified) and narrowed down, these methods worked well, but there are possibilities of misdiagnosis and inconsistency due to lack of data. The major limitation is the necessity for extensive data. The amount of required data is far larger than can be obtained routinely in the clinical process.

3.2 Mathematical Approach

A mathematical approach was also used in limited areas. The actual techniques are applied in detailed individual applications rather than as a whole program, such as calculating exact amounts of dosages of drugs in digitalis and acid-base/electrolyte disorders (10, 18). The mathematical formulations are based on blood level, body weight and renal function.

The best known program using this approach is a program to diagnose acid-based disorders developed at Boston Beth Israel Hospital (18). After collecting data, the abnormalities in the initial data evoke the chain of logic. The logic is similar to a flowchart approach consisting of different paths, but it involves complex mathematical relationships. All the questions are in the form of numeric values or boolean values.

The program generates a diagnosis and a treatment plan but there is no feed back in the system and every patient is considered as a new case. Also, the program does not have any conception of follow-up therapies and does not pursue specific etiologies and the possible relationships among diseases. This can be particularly bothersome in medical diagnoses.

The major strength of mathematical models is their efficiency and their ability to capture mathematically sound relationships. However, only few areas of medicine are adaptable to its rigid approach requiring correct identification of patient's physiological condition.

3.3 Flowchart Algorithm

The flowchart algorithm approach was also introduced and it is very different in medical reasoning from the previous ones. It is very

simple and easy to understand and is a direct implementation approach (10), consisting of the sequences of decisions which a physician can choose, but again, it has a severe drawback because there are several different pathways to the same conclusions and the different pathways have not always proved correct. Also, representing all possible pathways preferred by different physicians strictly by flowchart causes confusion. Furthermore, the flowchart does not include information about its own logical organization. There is no way to discipline for systematic revision or updating of the program, causing inconsistencies in modification and justification.

3.4 Decision Theory

Decision theory (2, 10, 18) is based on decision trees, and the decision making process is viewed as a sequence of steps in which the physician is navigated through a tree-shaped network of events and actions.

There are two types of nodes in the study of renal failure using this approach: decision nodes and chance nodes. At decision nodes, the physician must choose an action from a set of actions; and at chance nodes, definite actions are not taken but the physician gets the responses of the action taken.

In this approach, the expected value is the major decision analysis, and is calculated based on probabilities associated with the chance nodes by assigning numerical values to risks of therapies. Therefore, it is possible that the best possible outcome is not necessarily the desirable treatment, if the expected cost associated with the path far outweighs the other paths.

Perhaps the most serious problem with decision theory is assigning a numeric value to a human life or to the condition of health. In addition, there is no provision regarding the time courses of diseases, and it requires extensive data about conditional probabilities of symptoms and diseases.

3.5 Pattern Recognition

Around 1970, various pattern recognition methods (2, 8, 18) appeared to overcome the problem of limited amounts of statistically available data. The major characteristic of these methods was the use of well-chosen heuristics, which were able to summarize the large number of findings, that is, findings from patients would be transformed mathematically into some heuristic scores or weights, which then became the only tool for medical reasoning. Typically, this process consists of filtering the findings to extract a set of patient findings relevant to the clinical problem under consideration, and transformation of the set of findings into logical mathematical form. The domain specific knowledge is used to compare various patterns of association between findings and hypotheses.

The most common difficulties of this method are choosing the set of patterns in the first place and the fact that the whole system is not flexible enough to emulate a human expert's behavior.

3.6 Summary

In conclusion, all these methodologies are geared more toward performance than understanding (18). Medical consultation is a problem-solving process requiring extensive knowledge, experience and

common sense. Medical diagnosis can be viewed as judgmental (19) since it is dealing with uncertainty and incomplete knowledge.

In addition to judgmental knowledge, there are other forms of knowledge that can play an important role in the medical reasoning process. They are the underlying scientific theories and relationships (19) providing the foundation for a human expert in the decision making process. Unfortunately, these theories and relationships are often absent from computer programs.

Medical diagnosis using a computer is a symbolic reasoning process requiring knowledge acquisition, representation, explanation, and even self-modification. Analysis of a physician's reasoning process suggests that, as decisions become more complex, the physician's reasoning process becomes less algorithmic but more heuristic (18). At this stage, the importance of qualitative judgemental knowledge and the conditions for invoking the knowledge increases. Therefore, it is not surprising that AI methods were introduced to computer-based medical decision making.

CHAPTER 4

ARTIFICIAL INTELLIGENCE (AI) METHOD AND ITS GOALS

An AI method can be defined as a computer application that involves symbolic inference rather than strictly numerical computations or information storage and retrieval. The concept of intelligence is that the computer is capable of exhibiting behaviors which would be termed as intelligent if such behaviors were demonstrated by a person. Thus, an AI method consists of the modeling and representation of knowledge, reasoning, deduction, problem solving, heuristic search, and language.

Non-AI medical diagnostic programs are able to recognize and use causal and temporal relationships among diseases and to reason about the disease mechanisms at various levels of detail, but they are unable to evaluate the interactions, commonalities and differences among diseases.

AI methods were first introduced in the MYCIN, INTERNIST, CASNET and PIP systems. These systems are often called knowledge-based because they require extensive, highly structured medical knowledge.

Medical decisions are based on information about a patient's illness. Because there are considerable facts or findings which could be obtained, it needs to be focused on particular facts or findings. This focus is obtained by the set of hypotheses. Usually the process of diagnosis is concerned with using the available facts rather than acquiring additional data. Therefore, a typical AI method is composed of two parts: one part to form the hypotheses that explain the data (or

knowledge) known to the system, and one part to acquire new information from the program users to clarify the hypotheses under consideration.

However, some important questions arise: What is medically expert knowledge, how much knowledge is required, what criteria should be used to organize the knowledge, and how should the knowledge be applied? The answer to these questions will come only after careful study of real problem domains. One needs clear understanding of diseases, disease-causing organisms, and other notions that are related to the diseases.

Furthermore, an AI method in medical diagnosis involves interpreting available information about a patient's illness and acquiring new information if necessary and discriminating acquired information among various hypotheses. In addition, it must take consideration of cost, efficiency, urgency, and also the possibility of errors. Thus, the goals of an AI method can be described as follows (2, 10, 21):

- 1) To capture the medical expert's knowledge about specific medical inferences in a structured modular fashion; each module contains a small subset of medical concepts or facts. This provides programs to explain reasoning process categorically whenever necessity arises. The knowledge can be pre-compiled for interpretation of individual cases, hence leading to better performance.
- 2) To allow the separation of the domain-specific knowledge base from the reasoning and control mechanism used by consultation systems. This provides better modification of the knowledge base which may require changes for additional information as a result of medical research.

- 3) To develop logically powerful representations for describing medical structure and concepts that can justify and support the decisions being made.
- 4) To provide flexible programs that can interact with users in language and notations that are simple yet subtle enough to represent complex medical knowledge.

By incorporating the above descriptions, consultation programs will have characteristics closer to those of human experts in depth, scope, and flexibility of inferencing.

CHAPTER 5

HYPOTHESES IN MEDICAL REASONING

In order to achieve a computer-based medical diagnosis with capability comparable to that of a skilled physician, it is necessary to develop synthetic reasoning, which emulates those aspects of a skilled physician's behavior used in making medical decisions.

Hypotheses stand for major medical concepts used in inferencing, such as diagnosis, prognosis, and therapy. Some hypotheses are concerned with the direct etiology of illness, while some others include symptoms, pathophysiological and clinical states, courses of illness, and groups of clinical findings. Sometimes it is necessary to include a patient's environment, family history, age, sex, and race in the formulation of hypotheses. Therefore, hypotheses are expected to be changed after acquisition of each additional piece of information. Some hypotheses may be subconcepts of others (10), and thus inherit properties of parent concepts, and some hypotheses may have a cause-effect relationship implying priorities to be considered. Also, hypotheses can be merged into more parsimonious units.

During the course of diagnostic inferencing, hypotheses must be either confirmed or disconfirmed, and certain hypotheses become active according to the confirmation of findings. Once a hypothesis becomes active, it must be evaluated locally and globally after each additional piece of information. Based upon this process, a hypothesis may be accepted or rejected before a definite decision is made.

A consultation system must consider various treatments without overlooking their applicability, risk factors, benefit factors, side

effects, and effectiveness. Also to manage a patient with a lengthy illness, it has to plan carefully in terms of sequences to follow periodically after the initial treatment. The relationship between hypotheses and findings from the patient is the most important basis for deriving correct treatment. Thus, it is important to formulate hypotheses realistically with an awareness of the large scope of alternatives.

An hypothesis is, then, one of the most basic components for synthetic reasoning in a medical consultation system, providing links between findings and clinical decisions.

Methods for the formation, reduction, and synthesis of hypotheses have evolved over a number of system designs. Earlier medical diagnostic systems employed a simple heuristic method dealing with problem formation and solution sequentially (14). This method proved to be effective in sorting out the complexities and reaching a correct diagnosis in a narrowed specific medical domain. However, it became apparent that many aspects of existing systems' performance could be heightened if it were possible to deal with interrelationships among hypotheses simultaneously.

Computer-based consultation programs are strongly oriented towards the identification of unitary hypotheses. Searching through causal networks or hierarchies, structured on the basis of pathophysiological concepts, is one way to formulate unitary hypotheses to explain a clinical problem comprising a number of interrelated hypotheses. In other words, computer-based medical reasoning searches out unitary

hypotheses that can explain all the data found in a patient, then synthesizes the hypotheses by merging conjunctive hypotheses into more parsimonious units.

CHAPTER 6

A CASE STUDY (DERMAS)

6.1 DERMAS as a Task Environment for AI

A medical problem can be characterized in terms of an initial state (illness), an intermediate state (treatment state) and a goal state (health). An intermediate state is controlled by various drugs, diets, and surgical procedures, and, according to the effectiveness of above control operators, one state can be transformed into another.

Because a diagnosis is focused on the cause rather than its manifestations, while signs and symptoms are presenting at the initial state suggesting a cause or causes, in most cases, pathophysiological evidence is the most conclusive. To manage complex pathophysiological evidence, there is need for the formulation of a model, comprising of one or more hypothesized pathophysiological facts as a basis for clinical problem solving.

The clinical problems of infection caused by dermatophytes has been selected in this report for several reasons. First, the number of diseases caused by dermatophyte infection is relatively small and manageable. Second, because it is heavily dependent on laboratory observation rather than clinical observation for definite diagnosis, it is necessary to have an extensive pathological data base for a very small subset of dermatose problems. This is particularly important because computer programs are known to be more successful in narrow, constraining single areas of medicine with much underlying pathophysiologic understanding and in situations where decisions are based largely on laboratory data (2, 12).

6.2 Domain of DERMAS

This system is designed for the case study, and is named "DERMAS" since it deals with a small subset of dermatose problems and superficial fungal infections caused by dermatophytes and it is designed to be used by a clinician who is familiar with medical terms.

Dermatophyte belongs to a unique class of fungi which attach to the superficial layers of the epidermis in nails and hair and proliferate and live there. Dermatophytes grow only within Keratin, because there is a potent serum antifungal factor present in all individuals that protects living tissue against deep penetration of fungal elements (24, 27).

The species of dermatophytes can be categorized as geophilic, zoophilic, and anthrophilic. Geophilic dermatophytes inhabit the soil and frequently contaminate human skin. Zoophilic dermatophytes are primarily parasites on animals rather than man. Anthrophilic dermatophytes are those primarily parasitic on man, and man is known to be the host to thirteen anthrophilic dermatophyte species (24, 27). (Refer to Appendix for more information.)

There are three types of anthrophilic dermatophytes. They are: Trichophyton, which usually affects hair, skin, and nails; Microsporum, which usually affects only skin and hair; and Epidermophyton, which usually affects skin and nails (24, 26).

DERMAS is designed to deal with the dermatose caused by anthrophilic dermatophytes for the case study but it can be extended to handle a wider range of dermatose problems.

The most definite way to diagnose this type of skin problem is identification of the presence of dermatophytes. Therefore, microscopic

examination of skin scrapings, hair, and nails free from dirt, ointments, anti-fungal agents, and cosmetics is the most important part of diagnosis. For laboratory examination, KOH (potassium hydroxide) is used for mounting fluid and a 20% solution of KOH causes the keratin cells to swell and clear and reveal the fungus hyphae (24).

Dermatophytes have distinctive macro and micro aleuriospores. An aleuriospore is produced by terminal expansion of a hyphal branch followed by it's separation from the parental hyphae (24).

If the infected area is the hair, examination under ultraviolet light (Wood's light) can be very helpful to make a diagnosis or to follow the course of therapy. *Tinea Capitis* caused by *M. Audouinii* and *M. Ferrugineum* usually flouresce brightly in yellow-green while other types rarely if ever flouresce (23, 26).

Appearance of the culture produced in an artificial medium is also very important for the correct diagnosis. The most effective uses of culture observation are as early stages of growth or from the active edge of thallus and after the thallus has matured. Sabouraud's dextrose agar is used most commonly for routine cultivation of dermatophytes. Sabouraud's agar can be modified by adding yeast extract, without dextrose or by diluting it. Hyphae and spores grow on the media, and identification of the species of fungi is established by the gross appearance of the mycelia, the color of the substrate, and the microscopic appearance of the spores and the hyphae when a sample of the growth is placed on a slide. Also, if DTM (Dermatophyte Test Medium) is used, the problem of distinguishing dermatophyte from other fungi is

much simplified, since most bacteria and many common contaminants of dermatophyte cultures fail to grow on DTM or do not change the indicate color (24).

Once dermatophyte fungi's presence becomes positive, diagnosis of clinical disease is rather simple, because fungal diseases caused by dermatophytes are classified according to the locations of the infection such as, Tinea Corporis (Tinea of smooth skin), Tinea Capitis (Tinea of scalp and hair), Tinea Barbae (Tinea of beard), Onychomycosis (Tinea of nails), and so on (25).

A dermatophyte infection usually lasts a long time, requiring lengthy therapy, and the therapy often continues for a while even after the fungi disappears from the affected areas. It can be treated both topically and orally. The most common oral medication used in dermatophyte infection is Griseofulvin. Dermatophytes are particularly known for their sensitivity to griseofulvin, but in certain cases, griseofulvin therapy is not recommended at all. Generally, oral griseofulvin therapy should not be used to treat the feet and tinea of the toenails, because the recurrence rate after therapy is completed is very high, even after a year or more of successful therapy. Determining the correct daily dose and length of treatment with griseofulvin is important (23, 26).

6.3 Major Components of DERMAS

The basic components of DERMAS are distinctively divided into two categories based on their functions, and each category resides in a different storage: The first component in the consultation system residing in long-term storage consists of a knowledge base, an auxiliary

data base, a knowledge acquisition program, an inference mechanism and a control mechanism; and the second component is the consultation process residing in short-term storage.

The knowledge base contains a rich collection of medical data such as diseases, clinical states, physiological states, signs, symptoms, laboratory data, time courses of given illnesses, and rules to determine how closely a given patient's data might match the disease. The data in a knowledge base are entered through a knowledge acquisition program, and the knowledge acquisition program checks the validity of entering data before it passes data to the knowledge base.

An auxiliary data base is an extra data base containing a medical dictionary, which is needed because of a lack of standardization of medical terminology. It describes a set of complete lists of different terminologies used for the given medical term. Also it has all the patients' medical records containing previous visits, diagnoses, and other relevant facts.

The inference mechanism includes: a question-answer program to interact with users; a pattern matching program to scan incoming patient's data, so they will trigger hypotheses; a local evaluation by partitioning hypotheses; scoring validity; global evaluation to put together the result of several partitioned interpretations; an explanation program to trace and justify reasoning; and a pre-compiling program to focus the hypotheses under consideration.

The control mechanism supervises the interaction of various reasoning programs, determining how or when they should be invoked. Concrete task definitions are shaped by these invocations of the subprograms. It guides the computer in taking the data from a patient

and oversees the operation of various subprocesses, such as selection of questions. It generates hypotheses to be tested, accepted, or rejected, by deciding which aspects of the knowledge base will be used in the consultation process. The major goal of this mechanism is to reach the best possible formulation of individual cases efficiently. The formulation will be transmitted to the problem solver for the actual consultation process. Therefore, the control mechanism monitors the progress of the consultation system activities and revises the task definition on the basis of new insights obtained from the knowledge base. Also, it determines when and how the various conceptualizations of a diagnostic task are to be brought to the attention of the problem solver, and it decides when the goal has been satisfied.

The second component, consultation process, is based on individual cases. The amount of information in the short-term storage is variable depending on the complexity of a given case. It is the actual place where each patient's data enters and is processed to be diagnosed through interaction with the knowledge base, which is kept in the long-term storage. Each patient's data are provided by a clinician.

For the consultation process, a special structure "frame" is used which was originally proposed by Minsky (7). A frame is a data structure representing a certain situation and each frame is connected with several kinds of information. Some of this information is about how to use the frame, about sequences of happening, therefore one can expect what is going to happen next, and about what to do if these expectations fail. Each frame has property slots that must be filled by specific data. The differences between the frames can represent actions, cause-effect relations, or changes in concept.

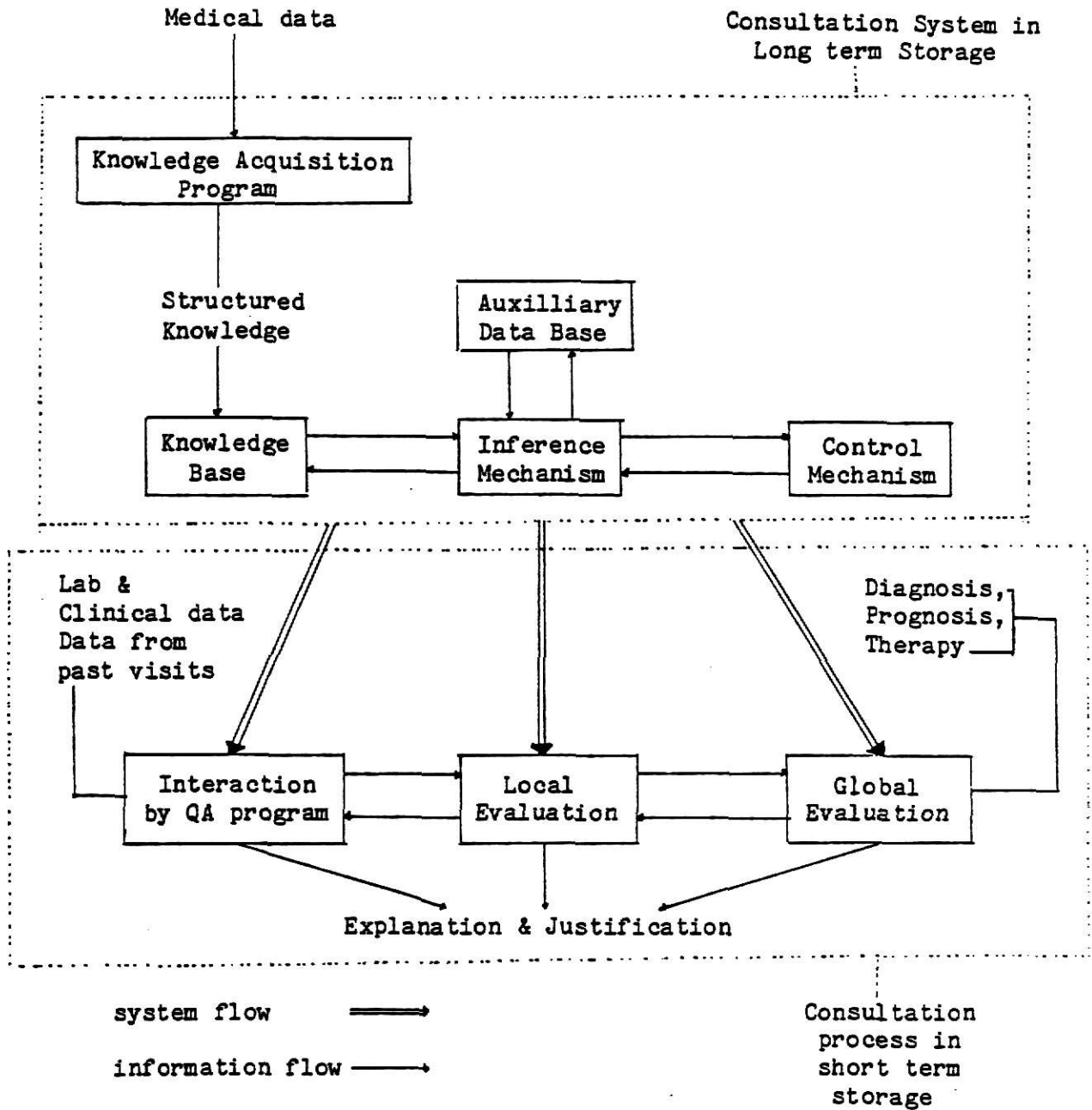
When an elementary hypothesis concerning a clinical disease is evoked, a frame is inserted into the short-term storage and the result of local evaluation of a patient's finding must fill the slots of different properties in the frame. Each frame contains a collection of related facts about a disease and clinical and pathological states. Therefore, a coherent picture of the patient is constructed in the form of a frame.

So far, the whole structure of DERMAS has been examined briefly and Diagram 1 shows interrelationships between different components. In the next section, the knowledge base will be discussed in detail.

6.4 Knowledge Base of DERMAS

The knowledge base is the heart of a computer-based diagnostic system. It serves as a repository for medical knowledge such as lists of all symptoms of a disease, lists of all characteristics of causing organisms, lists of possible organisms causing the particular disease, and so on. It grows automatically, incorporating new pieces of knowledge entered by clinicians from medical literature.

Therefore, a substantial part of creating a knowledge base is spent in knowledge acquisition. This system provides a high level of database editing mechanisms. The knowledge must be entered in machine-readable format but the editor handles all the bookkeeping involved in managing the rule base and error-checking system. As a piece of knowledge enters, syntactic validity and consistency with currently existing knowledge will be checked, and all stored knowledge can be easily displayed at request for evaluation.



Structure of DERMAS

Diagram 1

Generally speaking, knowledge consists of information, and information alone cannot be a part of knowledge unless properly structured. The problem of organization of a large amount of unconnected information is that, if it is not properly organized, things simply do not work, due to being out of control and out of focus. Therefore, some kind of formalisms are necessary for representation of knowledge.

One way to represent knowledge is using production rules. Functionally, production rules are decision rules, each consisting of a premise and an action. If the premise is true, the conclusion in the action part of the rule is drawn. In MYCIN (19), there are 400 production rules and each production rule represents a single chunk of knowledge, and all necessary conditions are included. This makes the knowledge base inherently modular and relatively easy to update. Individual rules can be added, deleted, or modified without drastically affecting the overall performance of the system.

For example, a typical production rule of MYCIN looks as follows (16):

```
if  the stain of the organism is grampos, and
    the morphology of the organism is coccus, and
    the growth confirmation of the organism is pairs, and
    the site of the culture is one of sputum lung-tissue,
then there is suggestive evidence (0.7) that the identity
    of organism is streptococcus-pneumoniae.
```

In this case, if not all of the conditions are satisfied, other rules will be tried to identify the organism. The numeric value in parentheses indicates the strength of a conclusion of a rule, the value being between 0 and 1. The larger the number, the greater the strength in the conclusion. In this case, the value (0.7) indicates the evidence is suggestive, but not absolutely certain.

Production rules are simple to understand and MYCIN (1, 2) has been showing considerable success in using production rules as a primary source of medical knowledge and inference engine of the system, represented in a form that is comprehensible to a clinician. The system uses a collection of rules to draw conclusions about the patient. If the answer still has not been discovered after trying all the relevant rules, the system asks the user for the relevant information which will establish the validity of the premise clause.

In ABEL (11, 19) and INTERNIST (2, 15), the medical knowledge is distributed throughout a hierarchy. The concept of a hierarchy provides the criteria to organize small collections of knowledge presented in the form of production rules. Concepts can thus be viewed as clusters of production rules. Production rules are extended to problem-solving situations. It is important to note that the concepts and not the rules provide the principle of a knowledge base (2, 15). Usually the most general concepts are placed at the top of the hierarchy and the most particular at the bottom. The structure serves the purpose of differentiating knowledge and integrating new knowledge. The top-most concept will first get control of the case, then the control will pass to an appropriate successor concept, and so on.

In the case of INTERNIST (2, 15), the diagnostic hierarchy is pre-determined. First the most top node establishes the concept of internist, then subnodes are established representing different diseases based on internal organs such as the heart, lung, liver, and so on. Each organ has subnodes consisting of particular clinical disease names. For instance, Hepatitis and Cholestasis are successors of the liver node. In INTERNIST, areas between nodes are not named.

In ABEL (11, 19), the lowest level of description consists of pathophysiological knowledge about diseases just as in INTERNIST, which successively aggregates into higher level concepts and relations, gradually moving the content of the description from physiological to syndromes of disease. The aggregation process allows a summary of the patient's illness at any given level, leading to the next more highly aggregated level. The summarization can be achieved by recognizing a central node and the chain of relations connected to the central node. This aggregate syndromic knowledge provides an overall perspective.

Using a strict hierarchy structure can be too restrictive in the representation of relationships between disease entities. Also it is annoying that there is no one right hierarchy of disease categories for purposes of assisting in the diagnostic reasoning process, because the value of any given hierarchic structure is dependent on the strong diagnostic relationships between diseases (2).

In the DERMAS system, semantic network structure is chosen to represent the knowledge base in the manner similar to its use in CASNET (2, 19). The semantic network is based on a diagraph, similar to that used in all networks, consisting of nodes and arcs. The nodes represent semantic objects and the arcs represent relationships among the nodes. It is possible to differentiate the nodes and the arcs to be more flexible and understandable. The knowledge base of DERMAS is a particular type of semantic network designed to describe diseases in causal terms, relate this description to clinical findings, and describe various categories at different levels.

The network consists of nodes and links primarily representing causes, effects, and explanations. There are two types of nodes:

elementary hypotheses, which are causes or explanations; and findings, which are basically manifestations, including history, symptoms, physical signs, and laboratory data. Elementary hypotheses are differentiated from findings, because they are subjected to evaluation during the consultation process. Some of the nodes are more specific, designating etiology, such as Onychomycosis, Tinea Pedis, Tinea Manus in Diagram 2. Therefore, an elementary hypothesis may or may not be the ultimate etiology.

Links between nodes specify the relationship of two nodes. There are several different types of links: Causal relations, represented by causing-organism and caused-by links; associational relations, represented by associated-with; explanatory relations, represented by is-a, may-complicated-by, and manifests; and finally grouping relations, represented by a-member-of. A set consists of several links of a-member-of.

A causal relation specifies the cause-effect relation between the cause and the effect nodes. An associational relation indicates that the presence of one node influences the expectation of the presence or absence of the other node. It suggests that two nodes are related but does not specify the reason for the correlation or association between the nodes. A grouping link expressed by a set and the links of a-member-of, representing the groups of nodes belonging to the same general category. When the category represents a group of organisms, two or more nodes may be present simultaneously in a disease. An explanation link is a description link, to clarify relations between two nodes. An is-a link goes from the more specific description to the general description. In the case of a node representing a disease, it

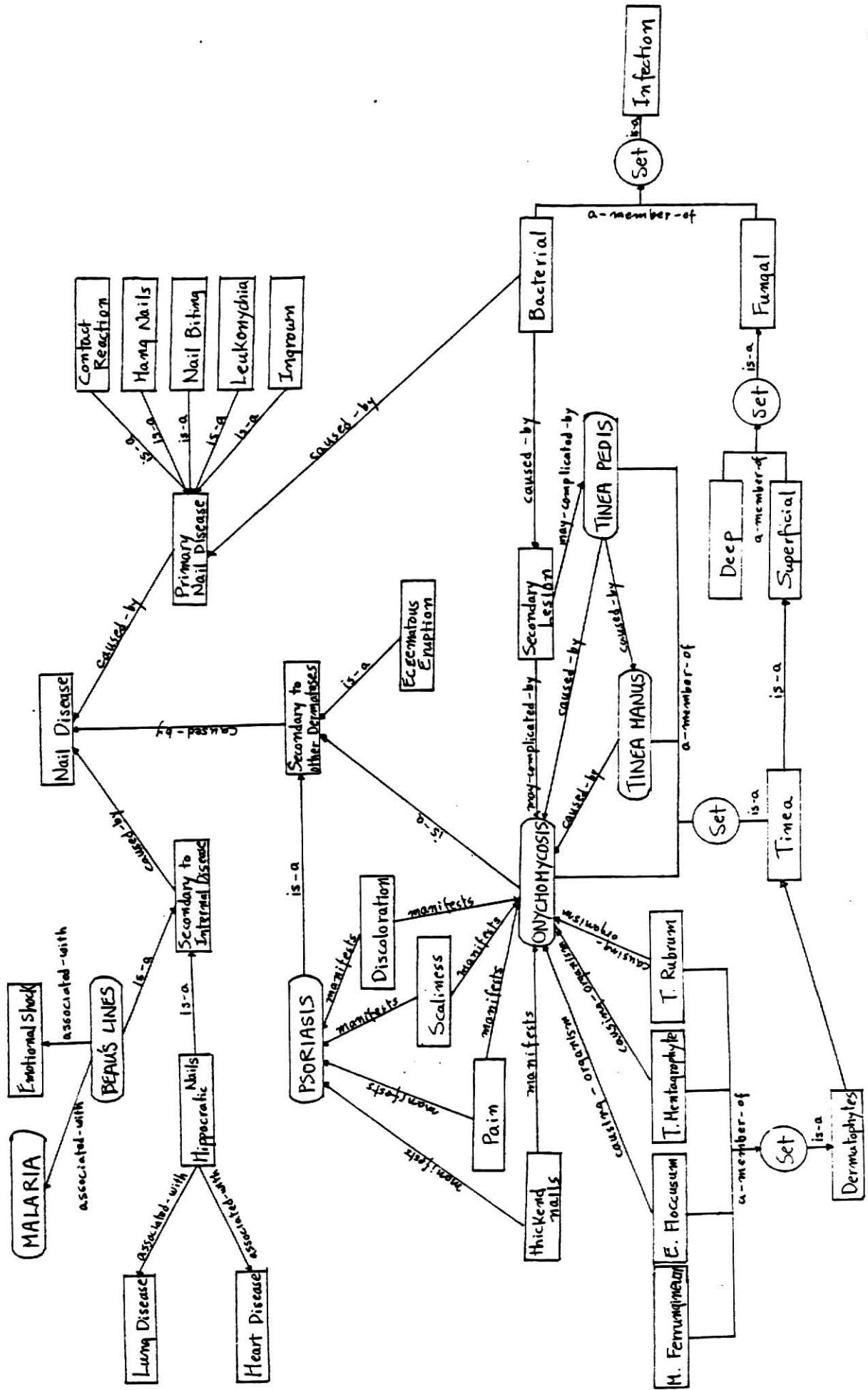


Diagram 1

has an associational list of manifestations to occur in that disease. There is also a set of disease categories organized around causing organisms.

Take, for example, the portion of the knowledge base network represented in Diagram 2, Onychomycosis is an elementary hypothesis. The symptoms of Onychomycosis include deformity of the nails, thickened nails, discoloration of nails and pain. Elementary hypotheses can sometimes themselves be considered symptoms, so that causal links, caused-by may connect them as in the cases of Tinea Manus caused by Tinea Pedis, Onychomycosis caused by Tinea Pedis, and bacterial infection leading to secondary legion which is a complication of Onychomycosis. Causing organisms of Onychomycosis are *T. Mentagrophyte*, *T. Rubrum*, and *E. Floccosum* which belong to the category of dermatophytes. In Diagram 2, an associational link is used to suggest that hypocratic nails may be associated with heart or lung diseases.

PIP also uses a network structure based on frames for its knowledge base. Each frame is a collection of data centered around diseases, clinical states and physiological states. Each frame has property slots to be filled. In the case of the clinical disease acute glomerulonephritis, symptoms, signs, and laboratory data such as anorexia and facial edema will be included in a typical finding slot; streptococcal-infection and sodium-retention, which are associated with causal relation will enter the slot of complementary relations to other frames. The other property slots include logical decision criteria, decision rules for excluding (must-not-have) and satisfying (is-sufficient) the fit of frame to the case at hand, storing in numerical numbers (Edema with severity = mass -> - 1.0), differential

diagnosis category (Recurring-Edema implies Nephrotic-syndrome), and so on. The frames are linked into a network of a variety of relationships; therefore, when input data are matched to an internal structure, a certain frame becomes active (2, 13).

In this section, the representation of the DERMAS knowledge base with its nodes and links has been examined in detail differentiating it from the other methods used for formulation of a knowledge base. In the next section, the logic of DERMAS will be discussed.

6.5 The Logic of DERMAS

The problem solving for this task will be performed in a topdown manner. Diagnostic logic employs a heuristic method for partitioning diagnostic inferences, which enables construction of a clinically relevant diagnostic model dynamically in the form of a frame for a given case, even in the presence of erroneous and misleading data.

The system tries to emulate a skilled physician in a real world situation by utilizing an extensive knowledge base and an auxiliary data base. Given a set of observations, a skilled physician generates tasks early in the patient encounter before definite data are available to ensure that his hypotheses are indeed appropriate ones. His rich network of knowledge enables him to develop his conceptualization of a clinical problem by working with what is initially available in general characters through a succession of refinement until specific disease entities are considered.

In DERMAS, a set of both clinical and pathological findings evokes hypotheses of symptoms, signs, and other syndromes, then of disease entities with which they are associated. Through this direct evocation,

certain findings may suggest more than one disease and one more strongly than another. Therefore, it is necessary to provide for some measure of strength of association between findings and the hypotheses they trigger. This evocation changes an elementary hypothesis from an inactive state to an active state.

Then hypotheses are generated in order to be evaluated locally, taking the new finding into account. During this process, the system does not ask questions about the status of other hypotheses or consider symptoms which are not relevant to the hypothesis being evaluated. Since the business of hypothesis formation and evaluation is an inexact process, it is necessary to provide some means for deciding among contending hypotheses. This suggests that the candidates to be scored on some basis and that criteria be established for deciding when the weight of evidence is sufficient. All these locally evaluated hypotheses will be parts of a frame for a later global evaluation.

The final and most important consideration is global evaluation; taking various elementary hypotheses and findings into a larger coherent structure, a frame. All frames are mutually exclusive, corresponding to coherent problem areas. This is essential in order to be able to deal with cases where more than one disease may be present and, therefore, more than one hypothesis of etiology correct.

The consultation process will be terminated when all the frames are evaluated, generating an explanation and justification of each frame, with their associated validity scores and treatment plans.

In the next section, each step of the consultation process will be examined in detail.

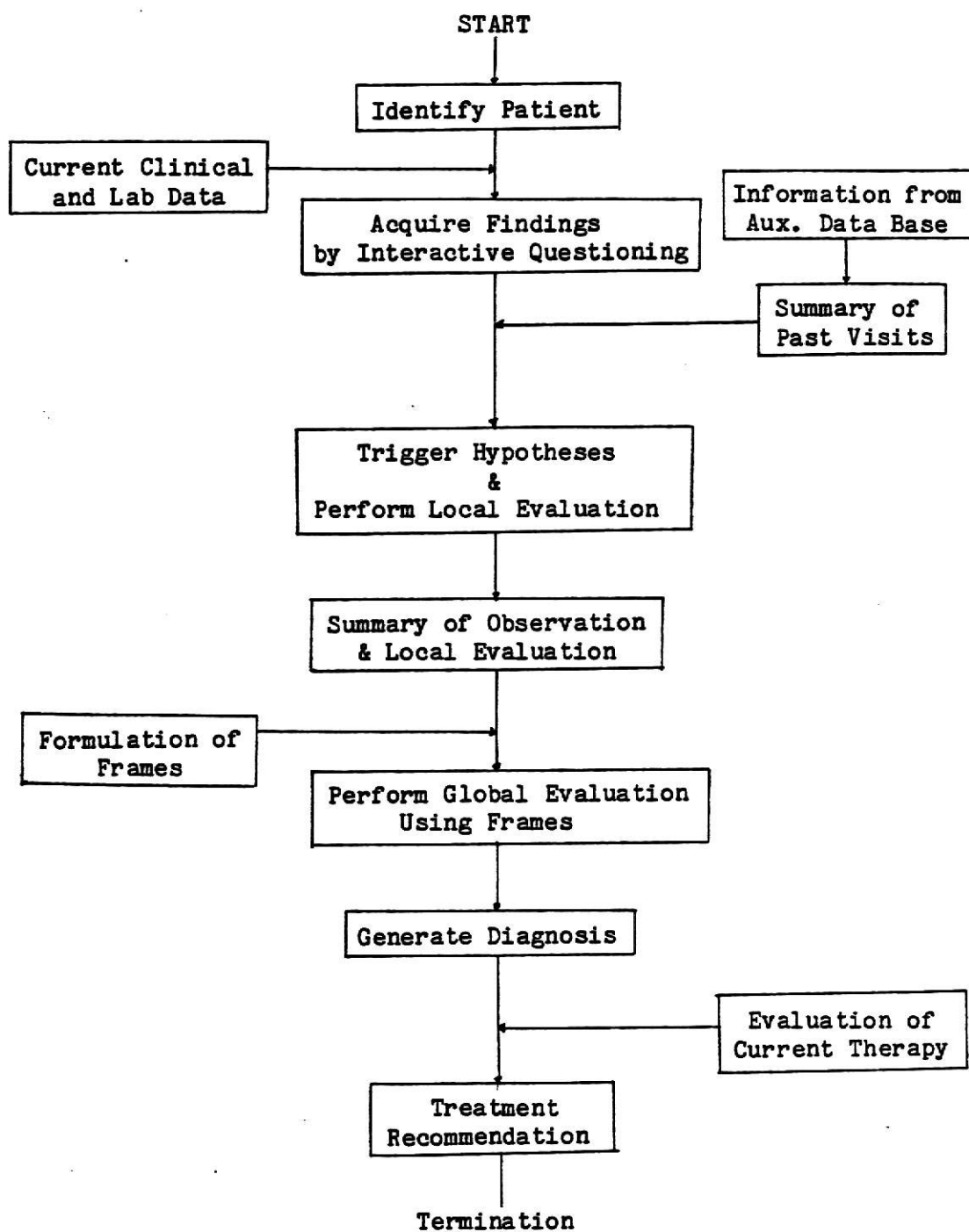
6.6 Consultation Process of DERMAS

In a computer-based diagnostic system, the problem is pre-determined, and the system's job is to select one or more diseases that best fit the facts of a case. In cases where more than one disease may be present, however, it is necessary to partition the set of diseases evoked by a given set of observations into different categories, each of which meets the problem criteria given above.

DERMAS employs the approach which focuses on one problem at a time using a frame with each successive problem dynamically evaluated by the clinical and laboratory data of the case developed up to that point. Each frame represents a single hypothesized clinical disease under consideration. When a frame is evaluated globally, according to the result of each property of a frame, a score will be assigned reflecting its fitness to the disease. Depending on which disease frame emerges as highest in score, the focus of attention may change from one frame to another, but one at a time. There is only one frame under consideration actively.

The program begins entering the clinical and laboratory data of a patient. The data can be deposited into the system by two different modes: user initiative and the system initiative. The system initiative mode makes it possible for that system to play a more active role by posing questions to the user. The questions are about relevant historical items, symptoms, signs, and laboratory data in multiple choice, yes or no, or short answer format.

The interaction between users and system is facilitated by the Question-Answer (QA) program. The QA program of DERMAS is very similar to the one in MYCIN (16). The QA program accepts natural language



Flowchart of Consultation Process

Diagram 3

questions about the system's knowledge of dermatophyte infection diseases and translates the question into a specifier to clarify hypotheses in the knowledge base related to the question, or vice versa. It also retrieves all the hypotheses if necessary.

The QA program includes the question selection function which analyses the available diagnostic questions and chooses the questions to optimize the number of questions related to the given case problem. When a user takes an initiative role, the QA program will take note of irrelevant, inappropriate, or missing questions which the system considers useful for the given case, and it will list out the questions to notify the user.

For example (26, 27), suppose a male patient, age 45, comes in with three fingers and two toenails infected with symptoms of discoloration, and surrounding skin shows mild redness and scaling. The infection's duration has been about 9 months. The patient's toenail infection causes mild pain when he wears certain tight-fitting shoes. Scaliness of the soles of the feet is also evident and mild bacterial infection is suggested. He doesn't have a history of nail infection but did have infection of feet in the past. Diagram 4 presents a portion of a typical dialog between user and program, for the patient with the nail infection described above.

The system knows about several different types of cases: A new entry to the system, a case already stored in the auxiliary data base to be retrieved for a review, or an old case that needs to be updated.

All these input data are subject to pattern matching with hypotheses in the knowledge base by an inference mechanism residing in the long-term storage, so they will trigger hypotheses. In other words,

Case No. 12

A Presenting Case: A man with nail infection

```
*Enter patient's name
  James Wilson
```

*Location of infection
finger and toe nails

*Indicate duration of infection

1. days	
2. weeks	
3. months	
4. years	3

```
*Is it
  1. first time?
  2. infrequent?
  3. occasional?
  4. frequent?      1
```

■Is patient using any medication? no

*Is there any discoloration of nails? yes

If yes, is it

1. severe?
2. mild?
3. moderate?

2

*Enter the identity of organism
T. Tonsurans

*Are there any other organisms under consideration?
yes

*If yes, list
T. Mentagrophyte

Some examples of dialog by QA program
(System initiation mode)

Diagram 4

a symptom is asserted and is scanned through the knowledge base, and it evokes all those elementary hypotheses which have been designated by the relationship links, "manifests" and "caused-by." For example, a symptom, discoloration of the nail, enters the system and after the symptom is scanned through the knowledge base in a heuristic manner, both elementary hypotheses containing Onychomycosis and Psoriasis will be activated because both of the two clinical diseases have the same symptom. The above example is shown in Diagram 1.

The hypotheses in active states are accepted hypotheses, and inactive hypotheses are rejected hypotheses. This process allows the diagnostic problem to construct the patient's differential diagnoses as generally as possible.

Each of these active hypothesis has an associated local evaluation function which produces a value representation of how likely the disease is to be present, given the data. Correlations between symptoms and diseases are the major function of local evaluation. Traditionally, to determine the strength of the correlation, consultation systems have used Bayesian conditional probability of a disease.

In DERMAS, a local scoring algorithm will be activated to determine its validity score, as in INTERNIST and PIP (10, 11, 19). In general, the presence of relevant symptoms will add to the validity score of an elementary hypothesis (in this case, a clinical disease) while the absence of symptoms will subtract from it. The validity score starts from zero and the range of the score is between -1 and +1.

For example, the presence of a thickened nail will add to the validity of Onychomycosis, while its absence will subtract. Later, during global evaluation, scores will be normalized by being divided by

their highest possible total score. The scoring algorithm also takes the sex and age of the patient into account. For example, patchy hair loss in a child rather than in an adult increases the score in diagnosis of Tinea Capitis.

Also a pathological hypothesis evaluation must be included as a local evaluation for identification of causing organisms, based on laboratory data. This is the single most important factor since identifying causing organisms provides the major conclusive evidence toward diagnoses in dermatology. For this, DERMAS employs a pre-compiled goal-directed method.

In the above patient, as an example, when all the manifestations evoke Onychomycosis, they also evoke the causing organism hypotheses directly connected to Onychomycosis by causing-organism links, such as T. Mentagrophyte, T. Rubrum, E. Floccosum, among the group of dermatophyte fungus, instead of activating the whole group of dermatophytes. This selective activation of hypotheses is one way to reduce the number of hypotheses actively being considered at any given time.

Throughout the consultation process, in response to certain user commands, the system offers summarization of the current and past lines of reasoning through an explanation program. Diagram 5 shows the summarization of the patients with nail infection after obtaining both clinical and laboratory findings. Some of the explanation available via the QA program accepts natural language questions dealing with any conclusions drawn during the consultation, or about the domain in general. The explanation program has the ability to examine any inferences that were made and to find out why others failed. This is

Case No. 12

A Presenting Case: A man with a nail infection

The case can be summarized as follows:

This is a middle-aged man, who has a nail infection, which is painful occasionally. It is the first time, and the condition has existed for 9 months. In the past, he did not have nail infections but did have infections of the feet. His nails show discoloration, brittleness, greater thickness than usual, and the surrounding skin is red with scaliness. He also has an infection in his left foot, and the presence of secondary lesion is reported.

A culture is drawn from debris of the nail plate and the skin scraping from the foot. The appearance of thallus are fluffy white with blood color on the underside as the culture matures, constantly developing a reddish pigmentation on a cornmeal dextrose agar.

Under the microscope, the fluffy thallus consists of long strands of hyphae with small, lateral, tear-shaped microaleuriospores. During sporulating, microaleuriospores are larger, clavate to round, and in small clusters along hyphae. Hyphae tend to break up into numerous arthrospores. Macroaleuriospores are rare. They appear to be less fluffy, slower in growth, heaped up and is less intense in pigmentation. Macroaleuriospores are narrow, long, and pencil-shaped, often developing directly at the ends of the hyphae. Potato dextrose agar is used as the testing medium. The organisms grow well in the absence of thiamine. Special nutritional requirements and perfect state are not known yet.

According to the above observations, the accepted hypothesis is *T. Rubrum* as causing organism.

An Example Generated by the Explanation Program
(Summarization of Patient of Diagram 4)

Diagram 5

accomplished through cross referencing criteria of different inference mechanisms, the knowledge base and the auxilliary base.

When the local evaluation of elementary hypotheses is completed, the system will generate one or more frames based on clinical diseases. Each frame is formulated dynamically by a control mechanism, and the properties of each frame are different and pre-determined based on clinical disease. The frame is the key structure for a global evaluation, and when there is more than one frame to be evaluated, they are mutually exclusive.

For a control strategy, DERMAS applies the approach of a binary choice task (True or False) as in INTERNIST and PIP (19). The task is based on the manipulation properties of the frame with scores. Each task has a choice between the presence or absence of a particular disease, meaning the disease either is or is not present; thus, it can not be both.

An ordinary Bayesian task formulation (18) requires the assumption that one and only one of the diseases in the differential list is present, while binary choice allows as many diseases as are known to the system. With DERMAS, the program is set up to consider the list of diseases, one disease at a time, shifting the focusing frame according to the validity score. According to the validity score, the focus of attention shifts from one to another, but the process does not stop because a hypothesis is confirmed to be the case sought.

This is different from INTERNIST and MYCIN (16, 19) in which programs form an interpretation of the patient's condition, which is essentially a list of possible diseases, ranked by a calculated estimate of likelihood. For example, INTERNIST employs a technique of

single-problem focus. It singles out the most highly scored hypothesis for consideration, from a list of diseases, which can account for the most important patient data. The diagnostic task consists of selecting one of the list by a process of elimination. If there are three hypotheses for the diagnostic task, and if two of them are ruled out as possibilities, then the third one is the case if the premises of the task are correct. Thus, it is possible to come to the correct diagnosis provided an appropriate differential diagnosis and considering remaining alternatives as its default. In this fashion, it is often possible to reach a conclusion concerning the remaining alternatives without the need of further investigation for definite evidences of confirmation.

When the local evaluation is over, all the results of the local evaluation are placed in the proper slots of the frames according to their category. The typical categories are: type of disease, list of findings, necessary conditions, local scores of findings, interrelationships with other diseases, differential diagnosis, list of unusual findings, list of common but missing findings in the given case, and so on. Diagram 6 shows the frame based on clinical disease of Onychomycosis with its properties after local evaluation of findings of the previously mentioned patient.

Since the DERMAS decision strategy is a binary task, the conclusion of the global evaluation will be confirmed or rejected based on the final validity score. The measurement of global evaluation is done by normalizing local finding scores, deciding how well each property is satisfied, and by additional normalizing of property scores. If the final score is above the pre-determined score, the hypothesis is considered confirmed. The single most important property is the

Case No. 12

Disease hypothesis: Onychomycosis

Is-a-type-of: Dermatophyte infection

Findings: pain
 discoloration of nail
 thickness of nail
 scaly and red surrounding skin
 fungal infection of feet
 presence of T. Rubrum

Necessary condition: Presence of Dermatophyte

Scoring:

clinical findings
 nail discoloration: 0.8
 thickness of nail: 0.5
 pain: 0.2
 infection of feet: 0.8

laboratory finding
 fluffy thallus: 0.5
 color: 0.7
 rareness of macroaleuriospores: 0.8
 tear-shaped hyphae: 0.7
 thiamine requirement: 0.8

causing organism
 T. Rubrum: 1.0
 T. Mentagrophyte: 0.6

Maybe-caused-by: T. Manus
 Occasional Tinea Pedis

Maybe-complicated-by: Bacterial Infection

Maybe-cause-of: Tinea Manus

Differential Diagnosis: nail injury
 Psoriasis of finger nails
 Psoriasis of toenails
 Moniliasis of fingernails

Unusual findings: None

A Typical Frame Based on a Disease Entity

Diagram 6

"necessary condition" property. If the case fails to satisfy this property, there is extra subtraction from the score; thus, the final validity score will not meet the necessary score to be confirmed.

For example, the case in Diagram 6 is a definite case of Onychomycosis, without any unusual findings, while satisfying all the necessary conditions. Therefore, the ultimate etiology, Onychomycosis becomes a confirmed hypothesis.

Although Onychomycosis is confirmed, since the DERMAS decision strategy is a binary task of all possible hypotheses, the other frames will continue to be evaluated. This is important to determine the co-presence of other diseases. In our example, as other frames are being evaluated, DERMAS also will confirm the hypothesis of Tinea Pedis which is also present in the same patient, but Psoriasis will be rejected.

Thus, the ultimate hypothesis either is a confirmed hypothesis or is a rejected hypothesis, with a single exception. The exception occurs when no single frame has a sufficient score to be a confirmed hypothesis. In this case, all the elementary hypotheses (frames are built upon them) will be suspended because there is not enough evidence to support the hypotheses. When this happens, the system will generate an explanation of the causes of failure and list necessary additional input data which will be needed for further investigation, at the user's command.

In our example case, there are confirmed hypotheses; therefore, the process of global evaluation will be terminated when all the frames are evaluated. At the time of termination of the global evaluation, at the user's request, the system will generate an explanation and

justification of each frame and it's diagnosis and treatment plan with possible recommendations. The typical diagnosis will be similar to the one illustrated in Diagram 7.

Since MYCIN was introduced, difficulties have been recognized in its backward chaining method (2, 22), (that is a method of going through a set of rules until a point is reached at which no further backup is possible and the user must be asked to supply information); to alleviate these difficulties, AI researchers have suggested different methods of inferencing.

A causal network (EXPERT, CASNET), frame (PIP), Hierarchy (INTERNIST) or combination of these (ABEL) are examples of alternative methods. CASNET (19) strictly uses a causal network throughout its medical reasoning. CASNET can be described as pathophysiological medical event driven model. An initial pathophysiological finding has a value associated with the level of confidence for the denial or confirmation of these states affected by the finding. The states are intermediate constructs that represent conditions or mechanisms that summarize results from many different findings. These states are linked to categories of diseases that are conceptually at the highest level of abstraction, summarizing both states and findings.

For any findings, a configuration of confirmed, denied, and undetermined states is formed. The starting states indicate the underlying basic mechanisms of disease. The most likely mechanisms for a given patient are identified by finding those starting states for which causal pathways may be generated which reach the largest number of states that are confirmed, without traversing a denied state. This process is repeated until all confirmed states are covered. Each

Case No. 12

Ultimate etiology: Onychomycosis

Diagnosis:

Onychomycosis is confirmed with the suspicion of Tinea manus and Tinea Pedis.

Prognosis:

The prognosis is very poor despite evulsion of the affected nails. A long period of treatment is necessary, and the progression will be slow, and a complete cure cannot be expected.

Causing Organism: T. Rubrum

Treatment:

- 1 Oral: Griseofulvin therapy is recommended for about 9 months
250 mg. q.i.d. after fatty meal
- 2 Topical: Onycho-phytex liquid, 15 ml. should be applied 2 to 4
times a day

Termination of treatment:

When there is no clinical evidence of infection and no evidence of fungi.

Recommendation:

Secondary lesion is due to bacterial infection, therefore, antibiotics should be considered. Also there is a high probability of Tinea Pedis and Tinea Manus. IF present, they need to be treated concurrently.

Generated Diagnosis and Treatment Plan

Diagram 7

starting state has a pointer to those classification tables of rules that are potentially applicable to the disease mechanisms related to that starting state.

The classification tables are logically ordered rules for combining confirmed and denied states in order to produce diagnostic, prognostic, and explanatory conclusions. They can be used to select different pathways through the causal net or to compose any configuration of states or pathways of states. Those rules that are logically satisfied will result in diagnostic conclusions for the patient (20). Although CASNET's performance has been satisfactory, defining diseases as paths in the causal network limits its primary reasoning mechanism to local evaluation of weights.

INTERNIST (15) uses a hierarchical approach, concentrating at higher levels of the disease hierarchy rather than at the terminal level nodes for problem solving. For example, jaundice, which is typically an observed symptom, is a strong clue to a general category of liver diseases, but provides no help in further partitioning within this sub-area. The existence of these specific patterns of association between common findings and higher level diseases has led to the inference that the subject is delineating the multi-problem structure of a clinical device. The major shortcoming of INTERNIST is that the program has a tendency to approach the analysis in complex cases inappropriately, thus wasting a lot of time (14, 19). This is due to its frequent reformulation of the task definition. The system has the tendency in problem solving to nearly always converge, eventually, on the appropriate conceptualization of a clinical problem (19). While it doesn't lead to false conclusions, it does prolong the sessions because

of its inability to perceive the multiplicity of problems in a case all at once.

Therefore, there is a need for a more flexible and general scheme of reasoning that will allow the introduction of empirical knowledge. In medical diagnosis, a physician uses many domain-specific concepts and rules of thumb to support his reasoning (19). Knowing how and when to apply such rules is very important for a correct diagnosis. This ability does not directly involve long chains of deductive reasoning used in MYCIN, but rather the ability to recognize and match patterns of findings and hypotheses. A diagnostic program must consider a case at various levels to integrate its overall understanding with its detailed knowledge.

DERMAS's extensive usage of frames is somewhat similar to PIP. The major difference between PIP and DERMAS is that frames are the primary structure of both the knowledge base and the consultation process in PIP while DERMAS makes these distinct. In PIP, the knowledge base is a network of frames; therefore, when input data need to be explored, a certain frame is pulled and the comparison between frame and patient data are made before decisions are made.

The DERMAS approach, a combination of network and frame, permits the overall system to be more flexible in medical reasoning. It's binary task in decision making allows it to deal with multi-problems with ease, and separation of different levels using different structures (network and frame) with their own criteria makes the whole system more consistent, since each structure double checks input data, and no contradiction between different structures is allowed.

6.7 Evaluation and Future Work of DERMAS

Like other existing medical consultation systems, the formalism of DERMAS does not capture all the aspects of a skilled physician. To deal with this problem, there is a need for new and more flexible strategies to be developed. The shortcoming of expression of complex theories of human problem-solving is also apparent. The system needs to be more dynamic in organization of many small problem-solving techniques into coherent strategy in larger and more complex environments.

Even though the binary choice approach offers advantages in differential diagnosis, the difficulty with the approach is that it fails to aggregate diagnostic possibilities into decision sets; instead, each considered diagnosis is evaluated as though independent of all other alternatives. This requires absolute criteria for decision making, as the problem solver is denied access to the powerful heuristics, discussed previously, that enables decisions to be rendered relative to a presumed decision set.

Since DERMAS is at the design stage, it is difficult to evaluate its acceptability, but it can be assumed that DERMAS will have to face the same problems as other existing systems, notably, biased opinions by medical people. The problem domain is so narrow and constrained that the program will have only limited application. The decision strategies may not be easily understood by physicians. Also, even after implementation of DERMAS, physicians may feel that it is too tedious and time-consuming to use on a practical basis.

The next task of DERMAS would be implementation. Building a system which can interactively assimilate the extensive knowledge and has a natural language capability based upon existing technology is no simple

task. Since DERMAS is dealing only with a small subset of medical problems, storing all the information needed can be done at a reasonable cost, particularly if an inexpensive mass storage device is also used.

Probably the most difficult aspect will be the problem of coding. A computer language can be used entirely for implementation, as the INTERLISP system used by MYCIN: LISP compiler, LISP editor, and Conversational LISP. However, the present system is slow and demands more computing power than most hospitals can afford. Also, it is proven that INTERLISP features have been useful as a development tool but are not all essential for the purposes of MYCIN (16).

Another option is combining languages depending upon a component's function and purpose. CASNET is implemented in FORTRAN for its speed, but the editing program is written in SNOBOL which is a language mainly used for text editing (19). Since speed, efficiency, and storage requirements are the major considerations of an implementation, careful study of existing technology is important before actual embarking upon implementation to heighten the above areas.

CHAPTER 7

PERFORMANCE AND ACCEPTABILITY OF AI CONSULTATION SYSTEMS

Although there has been steady improvement in representation of the knowledge base and symbolic reasoning for medical decision making by computers, there are important issues to be considered in system design, performance, and acceptability.

Even though it is difficult to evaluate a program in a simple manner, according to the study of the performances of existing medical AI programs, their performance is quite impressive. In the case of MYCIN (22), medical experts agreed with MYCIN's result of organism significance in 97.3% of the instances, 76.7% of the instances of MYCIN's choice toward organism identity were identical to the experts, and 90.9% of the results were acceptable. And in therapy selection, MYCIN selected an average of 1.5 drugs/patient (range 1-2) while the experts selected an average of 1.6 drugs/patient (range 1-3). MYCIN's overall performance was considered acceptable in 93% (14/15) of the cases. CASNET (19) had an acceptance rate of 95% in clinical proficiency and EXPERT (19) had a rate of 94% (131/139) in correct diagnoses.

Clearly, the above evaluation shows that medical programs are highly competent, but evaluations conducted on existing medical AI programs show there are definite institutional biases and differences of opinion. While MYCIN reflects the Stanford approach (2, 19, 22), some of the results are not considered quite as acceptable by outside experts. Another kind of bias is against formalism in medicine. For example (2), while 77% of clinicians judged the clinical proficiency of

CASNET to be at the very expert level, and 71% judged it to be applicable to research, only 45% judged it to be very important in health care. Therefore, it seems that physicians impose a considerably higher standard to computer programs than they demand of themselves.

CHAPTER 8

FUTURE WORK

In spite of the evidence showing that use of computers would be a major technological advance that would help in the solution of many medical problems, medical AI programs are still largely experimental.

New approaches to knowledge representation, language understanding, heuristic pattern recognition, and other symbolic reasoning problems still need to be improved. Even though some of recent AI programs show promise for handling concurrent diseases, assessing the time course of disease, and emulating the reasoning styles of experts, it is not clear how the system's attention moves from one hypothesis to another as soon as new data suggests another possibility (22). Also, better ways to define the program's termination criteria are needed, particularly when there is not enough evidence to support reasoning.

So far, most of diagnostic program's interaction between users and computer is done by abbreviated language. Better language facility which makes a program more acceptable and more pleasant for physicians to use is another important aspect that needs further work. Another way to make a medical program more acceptable is to increase its availability through facilitating its implementation on mini computers and micro processors.

In summary, while computers cannot replace the actual job of physicians, medical AI programs can be valuable to both physicians and patients with their extensive knowledge base for aiding physicians' decisions, and the trend toward AI medical consultation systems includes improvement of performance and acceptance of their goals.

BIBLIOGRAPHY

1. Bleich, Howard. "Computer-Based Consultation." The American Journal of Medicine, Vol. 53, Sept. 1972, pp. 285-291.
2. Cohen, Paul R. "Application-Oriented AI Research: Medicine." The Handbook of Artificial Intelligence, Vol. 2, William Kaufmann, Inc., 1982, pp. 177-222.
3. Davis, Randall; Bruce Buchanan, and Edward Shortliff. "Production Rules as a Representation for a Knowledge-Based Consultation Program." Artificial Intelligence, Vol. 8, pp. 15-45, North-Holland Publishing Co., 1977.
4. Davis, Randall. "Interactive Transfer of Expertise: Acquisition of New Inference Rules." Artificial Intelligence, Vol. 12, pp. 121-157, North-Holland Publishing Co., 1979.
5. Feigenbaum, Edward A. "Themes and Studies of Knowledge Engineering." Expert Systems in the Microelectronic Age, 1979, p. 3-25, Edinburgh University Press.
6. Fritz, Karl J.; Mary A. Polascik and Albert M. Potts. "Computer Assisted Diagnosis for Ophthalmology." Comput. Bio. Med., Vol. 8, 1978, pp. 223-228, Pergamon Press.
7. Gomez, Fernando and B. Chandrasekaran. "Knowledge Organization and Distribution for Medical Diagnosis." IEEE Transactions on Systems, Man, and Cybernetics, Vol. SMC-11, No. 1, Jan. 1981, pp. 34-42.
8. Gorry, G. Anthony; Jerome P. Kassirer, Alvin Essig and William B. Schwartz. "Decision Analysis as the Basis for Computer-Aided Management of Acute Renal Failure." The American Journal of Medicine, Vol. 55, Oct. 1973, pp. 473-484.
9. Komaroff, Anthony L. "The Variability and Inaccuracy of Medical Data." Proceedings of the IEEE, Vol. 67, No. 9, Sept. 1979, pp. 1196-1205.
10. Kulikowski, Casimir A. "Artificial Intelligence Methods and Systems for Medical Consultation." IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. PAMI-2, No. 5, Setp. 1980, pp. 464-476.
11. Patil, Ramesh S.; Peter Szolovits and William B. Schwartz. "Information Acquisition in Diagnosis." Proceedings of the National Conference on A.I., 1982, pp. 345-348.
12. Patil, Ramesh S.; Peter Szolovits and William B. Schwartz. "Casual Understanding of Patient Illness in Medical Diagnosis." Proceedings of the Seventh International Joint Conference on Artificial Intelligence, 1981, pp. 893-899.

13. Pauker, Stephen G.; G. Anthony Gorry, Jerome P. Kassirer and William B. Schwartz. "Towards the Simulation of Clinical Cognition." *The American Journal of Medicine*, Vol. 60, June 1976, pp. 981-996.
14. Pople, Harry E. and Decision Systems Laboratory. "The Formation of Composite Hypotheses in Diagnostic Problem Solving: An Exercise in Synthetic Reasoning." *Proceedings of the 5th International Joint Conference on A.I.*, 1976, pp. 1030-1037.
15. Pople, Harry E.; Jade D. Myers and Randolph A. Miller. "DIALOG: A Model of Diagnostic Logic for Internal Medicine." *Proceedings of the 4th International Joint Conference on A.I.*, 1975, pp. 848-854.
16. Shortliff, Edward H. "Computer-Based Medical Consultations: MYCIN." *Elsevier Computer Science Library*, 1976.
17. Shortliff, Edward H.; Stanton G. Axline, Bruce G. Buchanan, Thomas C. Merigan, and Stanley N. Cohen. "An Artificial Intelligence Program to Advise Physicians Regarding Antimicrobial Therapy." *Computers and Biomedical Research* 6, 1973, pp. 544-560.
18. Shortliff, Edward H.; Bruce G. Buchanan and Edward A. Feigenbaum. "Knowledge Engineering for Medical Decision Making: A Review of Computer-Based Clinical Decision Aids." *The Proceedings of the IEEE*, Vol. 67, No.9, Sept. 1979, pp. 1207-1223.
19. Szolovits, Peter (Ed.). "Artificial Intelligence in Medicine." *AAAS selected symposium 61*, Westview Press, Inc., 1982.
20. Weiss, Sholom M.; Casimir A. Kulikowski and Aran Safir. "Glaucoma Consultation by Computer." *Comput. Bio. Med.*, Vol. 8, 1978, pp. 25-40, Pergamon Press.
21. Weiss, Sholom M. and Casimir A. Kulikowski. "Expert: A System for Developing Consultation Models." *Proceedings of the 6th International Joint Conference on A.I.*, 1979, pp. 942-947.
22. Yu, Victor L.; Bruce G. Buchanan, Edward H. Shortliff, Sharon M. Wraith, Randall Davis, A. Carlisle Scott and Stanley N. Cohen. "Evaluating the Performance of a Computer-Based Consultant." *Computer Programs in Biomedicine* 9, 1979, pp. 95-102, Elsevier/North Holland Biomedical Press.
23. Conant, N.; D. Smith, R. Baker and J. Callaway. "Manual of Clinical Mycology." 3rd Ed., 1971, W. B. Saunders Co.
24. *Dorland's illustrated Medical Dictionary*, 24th Ed., 1965, W. B. Saunders, Co.
25. Hazen, E. L.; M. A. Gordon, and F. C. Reed. "Laboratory Identification of Pathogenic Fungi Simplified." 3rd Ed., 1970, Charles C. Thomas, Springfield, Ill.

26. Sauer, Gordon C. "Manual of Skin Diseases." 3rd Ed., 1973, J. B. Lippincott Co.
27. Stewart, William D.; Julius L. Danto, and Stuart Maddin. "Dermatology." 3rd Ed., The C. V. Mosby Co., 1974.

APPENDIXES

APPENDIX A

GLOSSARY OF MEDICAL AND TAXONOMIC TERMS

aleuriospores - an asexual spore formed by splitting off from the summit of a hyphae. See spore.

anorexia - a serious condition in which the patient loses his appetite and systematically takes but little food, so that he becomes greatly emaciated.

arthrospores - spores or segment produced by fragmentation of separate hyphae. Produced by dermatophytes in culture and in hair, nails and epidermal scales.

cholestasis - stoppage or suppression of the flow of bile.

coccus - a spherical bacterial cell, usually slightly less than 1 mm in diameter.

dermatophytes - fungi of the genera Microsporum, Trichophyton, and Epidermophyton, regardless of pathogenicity. A fungus parasite upon the skin.

DTM - Dermatophyte Test Medium, used for isolating dermatophytes, containing distilled water, phytone, dextrose, Agar Agar, phenol

red solution, 0.8 N HCl, Cycloheximide, Gentamicin sulfate and chlortetracycline HCl.

edema - The presence of abnormally large amounts of fluid in the intercellular tissue spaces of the body.

Epidermophyton - A genus defined by having multiseptate club-shaped macro-aleuriospores but no micro-aleuriospores. The single species is *E. floccosum*.

glomerulonephritis - A variety of nephritis characterized by inflammation of the capillary loops in the glomeruli of the kidney.

grampos - Retaining the stain or decolorized by alcohol in Gram's method of staining, a primary characteristic of certain micro-organisms.

hepatitis - Inflammation of the liver.

hyperbilirubinemia - an excess of bilirubin in the blood.

hyphae - A filamentous element or strand of fungus mycelium.

jaundice - A syndrome characterized by hyperbilirubinemia and deposition of bile pigment in the skin and mucous membranes with resulting yellow appearance of the patient.

macro-aleuriospore - An aleuriospore-type with predominantly two or more cells.

micro-aleuriospore - A type of aleuriospore consisting predominantly of a one cell, and of small size (usually less than 5 micro mm maximum axis).

Microsporium - A genus defined by having echinulate macro-aleuriospores and smooth micro-aleuriospores.

Monilia - A genus of yeast-like fungi, characterized by producing mycelia. The former name for a genus of fungi now called Candida.

mycelium - Filimentous mat or vegetative whole of fungus.

nephritis - Inflammation of the kidney: a diffuse progressive degenerative or proliferative lesion affecting in varying proportion the renal parenchyma and the renal vesicular system.

onychomycosis - A disease of nails of the fingers and toes caused by dermatophyte.

Papulosquamous - both papular and scaly.

perfect state - A state in which nuclear fusion associated with a sexual process occurs.

psoriasis - A chronic, recurrent papulosquamous dermatosis, the distinctive lesion being a silvery gray scaling papule or plaque.

Sabouraud's agar - Used for routine cultivation of dermatophytes, consist of distilled water, dextrose, peptone, and agar agar. This agar can be a modified addition of yeast extract, omission of dextrose and by dilution.

spore - the reproductive element of one of the lower organisms. An asexual spore produced by division within the walls of a mother cell.

sputum - Matter ejected from the lungs, bronchi and trachea, through the mouth.

Streptococcus - A genus of microorganisms of the tribe streptococceae.

thallus - The fungus organism considered as a branching continuum starting from a germinal spore or some initial center of growth.

Trichophyton - A genus defined by having smooth macro-aleuriospores (predominantly cylindrical) and smooth micro-aleuriospores.

tinea - A name applied to many different kinds of fungal infections of the skin, the specific type usually being designated by a modifying term.

APPENDIX B
ANTHROPHILIC DERMATOPHYTES

1. E. floccosum
 2. M. audouinii
 3. M. ferrugineum
 4. T. concentricum
 5. T. gourvilii
 6. T. megninii
 7. T. mentagrophyte
 8. T. rubrum
 9. T. schoenleinii
 10. T. tonsurans
 11. T. soudanense
 12. T. violaceum
 13. T. yaoundei
- E. - abbreviation of Epidermophyton
M. - abbreviation of Microsporum
T. - abbreviation of Trichophyton

EXPERT SYSTEMS IN MEDICAL DIAGNOSIS: A DESIGN STUDY IN
DERMATOPHYTE DISEASES

by

OH, KYUNG NA

B.S., San Jose California State University, 1975

AN ABSTRACT OF A MASTER'S REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Computer Science

KANSAS STATE UNIVERSITY
Manhattan, Kansas
1983

ABSTRACT

The development of knowledge-based Artificial Intelligence (AI) consultation systems has improved rapidly in recent years demonstrating near-expert performance. Most of these programs are focused on small areas of medicine and are based on a highly structured knowledge base. AI methods are powerful new approaches, usually accomplished by collaboration of both medical and computer experts, which may integrate two or more established techniques.

Each system builds a knowledge base first, then reaches a diagnosis through various decision strategies. Compared to earlier automated methods before AI methods were introduced, these systems attempt to simulate medical experts' conceptual structure and reasoning rather than to emphasize the optimality of decision making under general criteria such as information entropy. It is clear that AI consultation systems are under constant evolution to more sophisticated systems to handle the complexity of medical problems.

Although diagnosis, prognosis, and therapy are better than at any time in the past, there is greater dissatisfaction; therefore, the goal of an AI system is straightforward; to maximize the likelihood of correct decisions about diagnosis, prognosis, and therapy.

Also it is evident that there are needs for change in man-machine relationships and accessibility, to be accepted more readily by physicians. Improvement of man-machine relationships may be accomplished only by a closer relationship between medicine and computer fields.