WORKING PAPER
ALFRED P. SLOAN SCHOOL OF MANAGEMENT

COMPUTER ANALYSIS OF DIAGNOSTIC AND THERAPEUTIC STRATEGIES*

by

George

G. Anthony Gorry

492-70
October 1970

MASSACHUSETTS
INSTITUTE OF TECHNOLOGY
50 MEMORIAL DRIVE
CAMBRIDGE, MASSACHUSETTS 02139

COMPUTER ANALYSIS OF DIAGNOSTIC

AND THERAPEUTIC STRATEGIES*

by

George
G. Anthony Gorry

492-70
October 1970

*Presented at the Engineering Foundation Research Conference, held at Proctor Academy, Andover, New Hampshire, August 24-28, 1970.

Introduction

The purpose of this paper is to review some recent research on the use of a computer to solve diagnostic and treatment problems in medicine. A major result of this research has been the development of a computer program which can be employed as a consultant in a number of medical problem areas. Here the considerations which underlie the program are discussed. The basic functions of the program are outlined in a non-technical way, and an example of the use of the program is given. Then the results of the use of the program for several different medical problems are reviewed. Finally, an attempt is made to ascertain the potential of programs such as this in the delivery of appropriate medical care. Detailed reports on various aspects of this research are available in the literature (1, 2, 3, 4), and so the emphasis here will be on providing a general overview of the work and results obtained to date.

Modelling the Diagnostic and Treatment Problem

The use of digital computers in the selection of good diagnostic and treatment strategies has received increased attention in recent years. One reason for this interest is the general desire to improve the ability of the clinician to deal with the difficult problems which can arise in the management of a patient. A significant portion of the difficulty stems from the fact that the physician must sort out numerous possibilities and develop hypotheses about the state of health of the patient. The ability of the computer to store extremely large amounts of data, to enumerate many possibilities, and to perform complex logical operations suggests its potential value in this problem solving process. Before a computer can be used to

significant advantage in analyzing diagnostic and treatment strategies, however, precise procedures must be formulated for the means of inference required to deduce the clinical state of the patient from observed signs and symptoms, and a formalized capability must be developed for the prediction and assessment of possible therapeutic measures. In other words, the problem of performing diagnostic inference and weighting therapeutic strategies must be reduced to a problem of computation.

In order to better understand the requirements, a model of the diagnostic-treatment problem was formulated. The model is a mathematical one, but its principal characteristics can be discussed in terms of the way a physician deals with this problem. The model is discussed in these terms here because our interest is not in the technical details of this work. On the other hand, it should be noted that the model was not developed as a description of the way in which physicians operate. The purpose of the model is to permit the exploitation of the particular capabilities of a computer. Hence, in the next several paragraphs, when we are discussing the way in which a physician deals with the problem, we are using 'physician' instead of 'model' for convenience, and are not presenting a theory of human problem solving in the medical area. (The relationship of the model to the actual problem solving behavior of physicians is discussed in (6).)

In general, a doctor confronted with a potentially ill patient initially does not have sufficient information about the patient to decide on a diagnosis or on a therapeutic policy. The information he does have, however, in addition to his general medical knowledge and experience enables him to formulate some tentative hypotheses about the state of health of the patient. This opinion

will exert a considerable effect on the strategy which the doctor will employ in dealing with the patient. For convenience, let us say that the options available to the physician are _tests_ and _treatments_. By _test_ we mean any means for obtaining additional information about the patient ranging from simple questions to laboratory procedures to certain surgical procedures. He employs those tests which he expects to provide results of significant value in improving his current view of the patient's problem. The term _treatment_ will be used to refer to any means at the doctor's disposal to correct the health state of the patient. Treatments range from drugs to a variety of surgical procedures. The selection of an appropriate treatment for a given patient is strongly dependent on the correctness of the doctor's opinion about the patient's problem. The selection of the wrong treatment, for whatever reason, can have very serious consequences for the patient.

The value of the information obtained from a test is determined by the contribution which this information makes to improving the doctor's current view of the patient's problem and hence to reducing the risk of misdiagnosis with its associated cost. Hence the doctor is inclined to perform many tests. On the other hand, the tests available to him generally are not without some cost in terms of patient discomfort, time of skilled persons, money, etc. Thus there is a conflicting tendency to hold the number of diagnostic tests to a minimum.

As is discussed in (1), the doctor resolves these conflicting tendencies by performing _sequential diagnosis_. At a particular point in time, given his current view of the patient's problem, he can evaluate the choices available

to him. The basic choice is to employ a test to obtain more information, or to select a treatment in the hopes of curing the patient.

If he elects to cease testing and to make a diagnosis, the choice of a treatment implies a certain risk of mistreatment through a misdiagnosis. On the other hand, he can perform some test in the hopes of gaining additional information upon which to base his diagnosis and the resulting choice of treatment. In this case, he incurs the cost of the test selected. When the results of the test are known, and when they have been incorporated into his current view of the problem, he is faced with a decision problem of exactly the same form as the one which he has just solved. Thus a doctor can be thought of as solving a sequence of similar decision problems. At each stage of the process, he balances the cost of further testing against the expected reduction in the cost of treatment which the test results will permit. When, in the opinion of the physician, no test possesses the property that is expected to reduce the risk of treatment by an amount which exceeds its cost, he will cease testing, make a diagnosis, and treat the patient. Because the physician repeatedly updates his current view of the problem in keeping with the latest information available to him, the physician is able to develop effective diagnostic and therapeutic strategies.

Although this description of the manner in which a physician deals with diagnosis-treatment problems is simplified and somewhat artificial, it does emphasize the fundamental role that sequential decision making plays in the process. It seems clear that it will be necessary for a computer program to exploit an analogous capability (framed in terms suitable for a machine) in solving more general problems of this type.

The Development of the Computer Program

In this section, the basic components of a computer program to assess diagnostic and therapeutic strategies are discussed. These components directly reflect the view of the required problem solving process outlined in the preceding section. The discussion of the program is non-technical. Readers interested in the technical details are referred to (1) and (2).

The program has been implemented on the Project MAC computer system at M.I.T. This system simultaneously serves many users at remote terminals (typewriter-like devices). It is possible for the user to interact in an almost instantaneous fashion with a program in the system, engaging in a dialogue with that program. The program described here exploits this type of interaction with the user.

The program has three basic components. The first is called the information structure, and it constitutes the medical experience of the program. By changing the information structure, one can convert the program for use in a new problem area. This is the only part of the program which changes from one application to the next.

In addition to the diseases, signs, symptoms, tests, and treatments, the information structure contains two types of information: probabilities and utilities. The probabilities relate signs and symptoms to diseases. For example, one probability might be the conditional probability of red blood cell casts in the urine given that the patient has acute tubular necrosis. The program's understanding of various diseases is entirely in terms of the conditional probabilities which relate the variety of signs and symptoms and treatment consequences to those diseases.

The utilities of the tests, treatments, and treatment consequences are thought of as the subjective preferences of an expert. The utility of a test reflects the pain associated with the test, the cost of the test, the time of a skilled person required for the test, the risk of the test to the patient, etc. Similar factors are reflected in the utilities of the treatments and the treatment consequences. Utility can be thought of as the common denominator in terms of which all these diverse factors are measured. Utility assessment will be considered in more detail later in this paper. Here we only note that if the program is to make comparisons of risk and cost, a common scale must be established for seemingly diverse outcomes.

The second major segment of the program is called the inference function. Basically the task of the inference function is to establish the diagnostic significance of a particular test result. In a typical situation, a doctor confronted with a particular diagnostic problem must interpret the available evidence (observed signs and symptoms, etc.) in terms of his general medical experience. In other words, he employs a method of deduction which can accommodate both his general understanding of diseases and the individual instance represented by the patient before. The inference function of the program is the analogue of this capability in the physician. It uses probabilistic inference based on Bayes rule (1), (4), to obtain a probability distribution for the likelihood of each disease given the evidence to date and general medical experience. The latter is incorporated in the information structure of the program. It is this probability distribution, then, which constitutes the current view taken by the program of the given problem. This

view is updated whenever any new evidence is made available to the program. The updated probability distribution is one of the major factors which influence the strategy chosen by the program for dealing with a given patient.

The third component of the program is called the test/treatment selection function. Its purpose is to select at each stage in the problem solving process an appropriate test or treatment for use on the patient. By considering the probability distribution associated with the current view of the problem and the utilities of the various treatment consequences, this function can determine the best treatment to perform assuming that no further tests are to be used. The treatment chosen is the one which minimizes the expected risk, and it provides the standard used in evaluating the potential value of further testing.

In evaluating the potential usefulness of a particular test, the program considers the current view, the utilities of the various tests, and the likelihood of the possible test results. For each possible result of a test, the program can simulate the change in the current distribution which would occur if this result were obtained. The expected risk of treatment can be estimated for this new distribution. For each result of a test, the expected risk of treatment given the result is weighted by the likelihood of obtaining that result, and the sum of these products is added to the utility of the test to obtain the overall measure. A schematic representation of the factors considered in evaluating a test is presented in Figure 1. By analyzing <u>decision trees</u> such as the one shown, the program attempts to select the best test or treatment at each stage of the analysis.

SAMPLE DECISION TREE

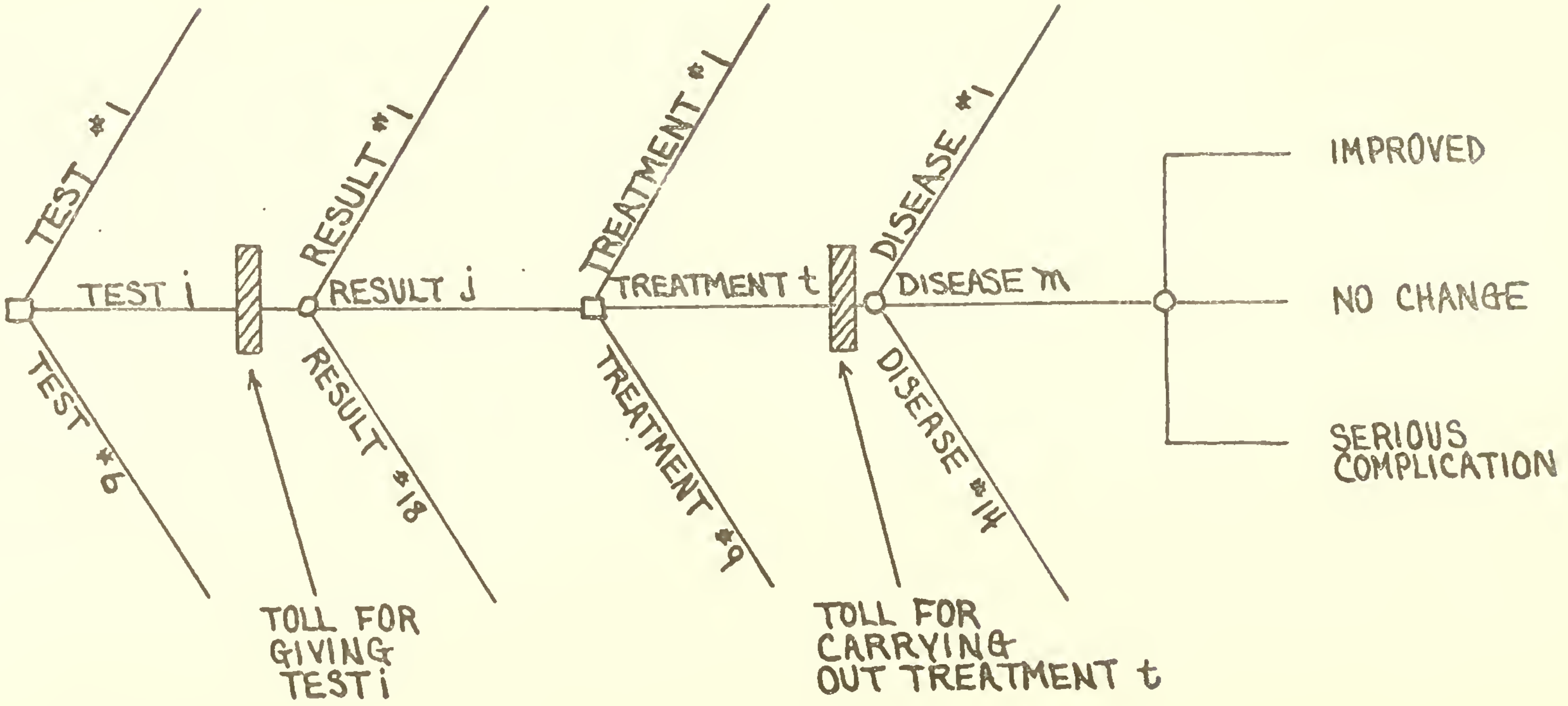

FIGURE 1

In Figure 2, an actual dialogue between a user and the program is presented. The problem being considered is the diagnosis of a case of congenital heart disease. At the outset of the discussion, the program is essentially passive, simply accumulating whatever evidence the user offers and using the inference function to update its current view of the problem. When the user has completed his initial description of the patient, the test/treatment function is invoked to determine the best diagnosis-treatment policy. In this case, no treatments were considered, and the problem was only one of diagnosis. The example, however, does give a basic impression of the use of the program.

Experience with the Program

The program has performed extremely well in the medical problems to which it has been applied. It was used as a diagnostic program (no treatments were considered) in two areas: primary bone tumors and congenital heart disease. The bone tumor problem involved the classification of patients into one of nine disease groups based on radiographic evidence and other physical findings. Studies made with the program indicated that it performed at the level of an experienced radiologist given the same evidence (1). The second medical problem considered was the diagnosis of congenital heart disease. Here there are 35 diseases in question. Again the results showed that the program performed at the level of an expert (1, 2). Both the bone tumor problem and the congenital heart disease problem had certain characteristics which made them unsatisfactory tests of the program's capability, however. First, in neither area were the costs of the tests sufficiently high to make

Figure 2

Example of a Session with the Program*

```
PLEASE TELL ME ABOUT THE PATIENT.
Well he is about 65 years old.
PLEASE CONTINUE.
No cyanosis.
O.K.  PLEASE GO ON.
What is your opinion of the 3 most probable diseases now.
       2 WITH PROBABILITY .68
      18 WITH PROBABILITY .10
      24 WITH PROBABILITY .04
ALL RIGHT.  GO ON PLEASE.
Name disease 4 for me.
ATRIAL SEPTAL DEFECT WITH PULMONARY HYPERTENSION.
PLEASE CONTINUE.
I meant the name of disease 24.
VENTRICULAR SEPTAL DEFECT WITH PULMONARY FLOW LE. 1.4 SYSTEMIC FLOW
O.K.  PLEASE GO ON.
What more do you need to know.
CHECK THE HILAR VESSELS.  ARE THEY INCREASED OR DECREASED.
That information is not available.
LOOK FOR A LEFT SECOND SYSTOLIC MURMUR.
What is your current view.
MY CURRENT VIEW IS
     DISEASE     PROB
           2      .68
          18      .10
          24      .04
          25      .04
          34      .03
           5      .03
          12      .02
NOW LET'S GO ON.
LOOK FOR A LEFT SECOND SYSTOLIC MURMUR.
Say there is a left second systolic murmur.
I SEE.
IS THE AXIS IN THE ECG RIGHT, LEFT, OR NORMAL.
What is the probability of atrial septal defect now.
IT IS 85.
ALL RIGHT.  GO ON PLEASE.
IS THE AXIS IN THE ECG RIGHT, LEFT, OR NORMAL.
Left.
ALL RIGHT.
CHECK FOR A LEFT FOURTH SYSTOLIC MURMUR.
What are the 3 most probable diseases now.
       5 WITH PROBABILITY .43
       2 WITH PROBABILITY .39
      18 WITH PROBABILITY .05
PLEASE CONTINUE.
CHECK FOR A LEFT FOURTH SYSTOLIC MURMUR.
What is the medical name of disease 5.
ATRIO-VENTRICULAR COMMUNIS.
O.K. PLEASE GO ON.
CHECK FOR A LEFT FOURTH SYSTOLIC MURMUR.
Say there is one.
O.K.
IS THERE AN ACCENTUATED OR DIMINISHED P2.
What is the probability of disease 5 now.
IT IS 71.
ALL RIGHT. GO ON PLEASE.
IS THERE AN ACCENTUATED OR DIMINISHED P2.
Not available.
I RECOMMEND A DIAGNOSIS OF DISEASE 5.
What are the 3 most probable diseases.
       5 WITH PROBABILITY .71
       2 WITH PROBABILITY .23
      24 WITH PROBABILITY .01
I'm finished.  Thanks for your help.
I HOPE THAT I HAVE BEEN OF SERVICE.  GOODBYE.
```

*User inputs in lower case; program responds in capitals.

the sequential aspect of the program particularly important. Because tests were so cheap, it mattered little how many were performed or in what order. Second, the treatment problem was not considered, because doctors to help define these treatments and their consequences were not available.

Because of these difficulties, a third medical area was considered, the diagnosis and treatment of acute renal failure. The management of the acute renal failure syndrome is an important medical problem. Although the incidence of the problem is relatively small in the context of all disease treatment problems, the potential risk to the patient's life is sufficiently great to give the problem special significance. In dealing with this problem, the physician must account for this risk factor as well as the possibility that the tests which he could employ to gain further information about the patient can contribute to the medical problem if improperly used.

The definition of the acute renal failure problem used in this study included fifteen diseases. As in the previous problems of bone tumors and congenital heart disease, the information structure for the program included the relevant probabilities. In this case, however, no attempt was made to obtain these probabilities from an analysis of historical data. Rather, the opinion of an expert was used in establishing each probability. In addition, special attention was paid to the assessment of the required utilities. Again the opinion of an expert renal specialist served as the basis for these numbers.

The precise manner in which these judgments were obtained from the expert and the way in which they were converted to utilities is discussed in (5). Here we want to briefly outline the procedure. The renal expert was

given a series of hypothetical decision problems. Each problem required him to make a choice between a particular event for certain (such as curing the patient by performing a certain operation) and accepting a chance in a lottery. If he chose the lottery, a given event would be chosen for him with probability 'p', and some other event would be chosen with probability '1-p'. Before making his choice, the expert is told exactly what the two events in the lottery are and what the value of 'p' is. With the theory discussed in (5), a series of these decision problems can be used to establish the utilities of tests, treatments, and consequences required by the program.

With the information structure for the renal failure problem developed in this way, the program duplicated the diagnostic-treatment decisions of expert renal specialists in over 90 percent of the cases tested. Furthermore, when the information structures from two experts were used, the program agreed more closely with the expert whose judgments it was using than did the other expert. This success suggests that the computer program (and the underlying model) is not only a convenient structure for theoretically describing diagnostic and treatment procedures, but that it is a practical way of analyzing such decision problems.

Conclusions

The research which has been reviewed in this paper provides strong evidence that it is possible to formalize procedures for use in computer-aided assessment of diagnostic and therapeutic strategies. The sequential mode of the program and the basic inference and evaluation functions seem to be fundamental to its success. Not only has the program proved successful in

diagnostic problems, but it has also been effective in a treatment problem in which there are potentially severe risks. Because the data base for the program in the renal failure problem was constructed entirely from an expert's subjective assessments of probabilities and utilities, the data collection problems were drastically reduced. In addition, no significant reduction in the performance of the program was noted.

Although more extensive tests of the program are required before extensive conclusions about its value can be drawn, the initial results are very promising. To the extent that the judgments of an expert about probability, risk, and pain can be captured in a program, new possibilities for the distribution of medical knowledge are realized. Thus the thrust of this research should be seen not as an attempt to eliminate doctors from decision making, but rather to make some portion of the judgmental expertise of the best specialists in a given problem more generally available to the medical community. To this end, programs such as the one described here merit continued investigation and development.

REFERENCES

(1) G. Anthony Gorry, "A System for Computer-Aided Diagnosis," Sloan School of Management Doctoral Dissertation, M.I.T., September 1967; also Project MAC Technical Report, MAC-44, 1967.

(2) G. Anthony Gorry, "Strategies for Computer-Aided Diagnosis," Mathematical Biosciences, 2, 1968, pp. 293-318.

(3) G. Anthony Gorry and G. Octo Barnett, "Sequential Diagnosis by Computer," Journal of the American Medical Association, Vol. 205, September 16, 1968, pp. 849-854.

(4) G. Anthony Gorry and G. Octo Barnett, "Experience with a Model of Sequential Diagnosis," Computers and Biomedical Research, Vol. 1, No. 5, pp. 490-507.

(5) Norman E. Betaque and G. Anthony Gorry, "Automated Judgmental Decision-Making," Sloan School of Management Working Paper 416-69, September, 1969; presented at 36th National Operations Research Society of America Meeting, Miami, Florida, November, 1969.

(6) G. Anthony Gorry, "Modelling the Diagnostic Process," Sloan School of Management Working Paper No. 370-69, M.I.T., February 20, 1969. To appear in The Journal of Medical Education.