WORKING PAPER 56

SOME ASPECTS OF MEDICAL DIAGNOSIS

by

GERALD JAY SUSSMAN

Massachusetts Institute of Technology

Artificial Intelligence Laboratory

December, 1973

Abstract

Since mid July Steve Pauker, Jerome Kassirer, and I (Gerald Jay Sussman) have been observing the diagnostic process of expert physicians with the goal of abstracting the underlying procedures being followed. One purpose of this position paper is to summarize our preliminary conclusions. I will attempt to pinpoint those aspects of the process we feel we understand, and where we are confused or unsure. I will also attempt to indicate some possible theoretical underpinnings of our ideas. Finally, I will propose what I consider to be a coherent research protocol for the development of these ideas.

Some Aspects of Medical Diagnosis

Since mid July Steve Pauker, Jerome Kassirer, and I (Gerald Jay
Sussman) have been observing the diagnostic process of expert physicians
with the goal of abstracting the underlying procedures being followed.
One purpose of this position paper is to summarize our preliminary
conclusions. I will attempt to pinpoint those aspects of the process we
feel we understand, and where we are confused or unsure. I will also
attempt to indicate some possible theoretical underpinnings of our ideas.
Finally, I will propose what I consider to be a coherent research
protocol for the development of these ideas.

Specifically, we have observed the behavior of several physicians,
of varying degrees of expertise, when presented with an unknown CPC. The
CPC was read to the subject, one fact at a time, and he was encouraged to
introspect verbally on his state of mind. Informal protocols were taken
and the subject was encouraged to participate in the analysis. Thus, the
subject picked up the language and beliefs of the experimenters and was
usually enthusiastic about the application of the budding theory. This
is perhaps a terrible way to do psychological research as the subject's
behavior is modified by the experiment; but we are more interested in a
good way to perform diagnosis than in precisely how any particular
physician operates. In any case, the subjects felt that, rather than
their methods being modified by the experiment, they had acquired a
language for describing the previously indescribable processes going on
in their minds.

Why did we feel that understanding the incremental processing of a CPC was a good place to start in understanding the cognitive processes of a physician? I believe that the CPC problem factors the diagnosis problem into some more manageable subproblems. While it focuses on the deep problem of the interpretation of and evaluation of evidence, it factors off the problem of what information to gather next, and its associated cost-benefit analysis. Besides, understanding the processing of a CPC would make it possible to produce a diagnostic consultant, a program which could accept CPC-like presentations and produce diagnoses. Indeed, much of what Steve Pauker does on "rounds" is very similar to CPC processing. One of the disadvantages of using CPC cases is the appearance of "CPC manship" -- the subject knows that CPC cases illustrate unusual or trick diagnosis and so he is constantly looking out for the trick. Another problem of CPC protocol analysis is the separation of thoughts which would normally occur to the subject from those which are the result of interrogation. We have noticed that physicians, besides having strategies for diagnosis and management of medical problems, have strategies, apparently somewhat independent, for explanation and exposition of decisions they have made.

II Some observations we have made:

A striking observation we have made is that the most expert physicians jump directly to a diagnostic hypothesis on a very small amount of information -- usually not much more than age, sex, and presenting symptom. Of course, such a diagnosis is very insecure; subsequent

information can easily refute it. In any case, as soon as the diagnostic hypothesis is made, subsequent data is processed in the light of that hypothesis. A new fact may tend to confirm, refute, or be indifferent with respect to the hypothesis. The diagnostic hypothesis is a framework of expectations. Thus, we have heard doctors react to new facts with such phrases as: "I expected that.", "(It is) consistent with my assumption.", "I did not expect that. Perhaps she really has lupus nephritis.", "That clinches it.", "This new fact is making me very unhappy with my diagnosis.", "It doesn't really fit in. Perhaps his congestive heart failure is complicated by a pulmonary embolus." Among the most important reactions are ones of the form: "This does not really fit in. Perhaps he has.....". In this case we find that if the new fact does not find an appropriate place in the expectation framework it may cause the old hypothesis to be scrapped in favor of another one. In this way, the physician jumps from hypothesis to hypothesis until he finds one which accounts for all of the abnormal findings, and which the patient-specific data matches reasonably well.

I feel that this is a surprising observation. A priori I expected that inexperienced diagnosticians would act this way and that experienced expert ones would reserve judgment until all of the available facts were in, using each fact to precisely narrow the field of possible choices among diseases using much physiological knowledge. A remarkable incident went far in showing me why this is not the case:

One day, on rounds, I watched a rather inexperienced physician present a case to Steve Pauker. His presentation went something like

this:

> "Mrs. X is a 57 year old mother of two children. Her husand is
> an executive. She has no fever or ankle swelling. She has no
> history of cardiovascular disease but her parents both died of
> heart trouble. She has a physiologically split S2 and a 2+
> systolic murmur heard best at the apex. Her neck veins are not
> elevated and she has a left carotid bruit. Her blood pressure
> is..."

At this point Steve began to fidget and soon he interrupted with, "So
what does she have? A toothache? An upset stomach? Why did she come to
the hospital? Please begin again." On the second try, the neophyte
began again with: "Mrs. X is a 57 year old woman who presented with
chest pain." This time Steve was happy with the presentation. Later on,
Steve and I analyzed the situation. He said that without the presenting
symptom to organize himself around, the facts were not useful; he could
not even remember them. The presenting symptom was necessary to provide
a framework for organizing the data. The 57 year old woman with chest
pain brings to mind several possible diagnostic hypothesis, each of which
provides a set of slots for the data items to be fit into. Without such
a set of slots the items just get lost in a mass of rubble. It seems
that it is easier for a physician to handle a coherent story than a
disorganized data set. Even discrepancies are easier. It is easier to
describe Mrs. X as having something that looks like Y except for symptom
Z (even if symptom Z rules out Y) than as a list of unbound symptoms.

III   The Frame Theory

Our observations indicate that the diagnostic knowledge of an expert physician is tightly organized. The resulting structure is (probably) optimized to make the best use of the limited computational resources of the physician. Just what is this structure? The structure of the organization of diagnostic knowledge that we propose is a development of Minsky's proposed theory of _frames_. In this theory the frame is the unit of recognition knowledge. A frame is (in the case of medicine) a chunk of the clinical picture of a patient. For any given patient a frame may be true, false, or somewhere in between. A frame may be anything from a well defined ultimate etiology of the symptoms, such as Acute (Post Streptococcal) Glomerulonephritis. A frame may be a physiological situation, such as Glomerulitis or Sodium Retention. A frame might also be a syndrome -- a less clearly related set of symptoms -- such as Nephrotic Syndrome.

Frames may be related to each other in a variety of ways. Acute Glomerulonephritis (AGN), for example, is one of the diseases classified as a Glomerulitis. A patient with AGN is likely to manifest Sodium Retention and may develop Nephrotic Syndrome as a complication. If we think a patient has AGN except that there is no evidence of a preceding strep infection and we notice a cardiac murmur, we should check if he has Sub-acute Bacterial Endocarditis (SBE).

Just what is a frame? We have seen that frames may be related in many ways. A frame is a chunk of a clinical pattern or picture along with a way of determining how well it fits a particular situation, and

pointers I call _differential_ _pointers_ by which particular types of misfit indicate (or suggest) other frames that might fit better. Each frame contains, among other structures, a _prototype_, or typical case, against which the data is matched. The prototype consists of a set of _slots_ -- named, relevant features of the prototype, the parts of the prototype -- against which the findings of the particular case are matched, and a set of relationships between the slots which make up the prototype. Consider, for example, the frame Glomerulitis. A person with Glomerulitis typically has Hematuria and Proteinuria (perhaps because his glomeruli have developed pores big enough to pass protein molecules and sometimes red blood cells.) The protein may congeal in the tubules (entrapping red blood cells) causing the important finding of Red Blood Cell Casts. If the Hematuria is Gross then only a small amount of Proteinuria is not good evidence for Glomerulitis because such protein could be accounted for as just the standard serum protein appearing in the urine from a bleeding wound. Thus the Proteinuria must be enough so that it cannot be simply explained with the Hematuria as a result of internal bleeding. Perhaps the essence of this prototype can be graphically expressed (see figure). Note that the slots Proteinuria, Hematuria, etc. of Glomerulitis are themselves frames. We find that such a frame as Proteinuria can have properties such as Light which can be interrogated and logically manipulated.

It is difficult to continue to describe this frame theory without describing how frames are used for diagnosis, how the matching process proceeds. Associated with each node in the frame network is a _status_
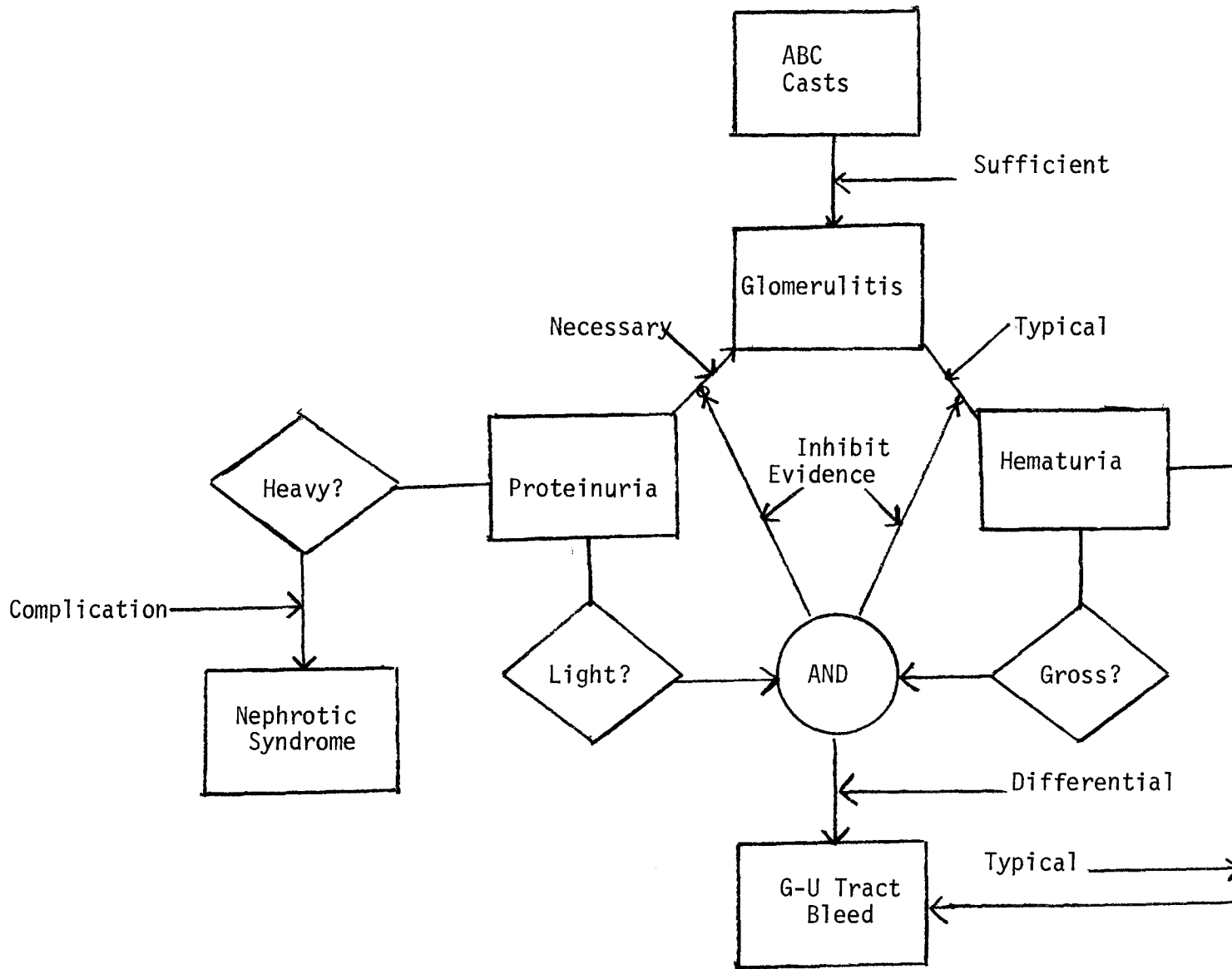
which describes that node's opinion of its relevance to the case being described. Initially all nodes are dormant. As facts of the case being described are read, they are bound to nodes in the frame network. Thus, if we read that patient X has gross hematuria, the node (or frame) Hematuria would become active and acquire the property Gross. Because Hematuria is surely true--it is bound to raw data--it becomes Happy. Activation then proceeds to other nodes connected to Hematuria. Glomerulitis, for example, is triggered; it goes from dormant to active because Hematuria is designated as a trigger slot in Glomerulitis. (This is depicted in the diagram as a dot on an arrow head. Although in this case all slots are triggers, this is not in general true: Though Glomerulitis triggers AGN, Sodium Retention does not.) Each frame so triggered runs a local evaluation function (not shown in the diagram -- each node has one associated with it) which determines if it is Happy (or how happy it is). Glomerulitis, for example, is only Happy if both Hematuria and Proteinuria are present or if there are RBC Casts present. A node, once triggered, remains active and thus watches for changes in its slots, recomputing its happiness for any change in the happiness of its slots whether or not they are trigger slots. If a frame becomes Happy it puts itself on a list of Happy frames and informs all of the frames of which it is a slot, that it has become a finding. At some point we ask the physician for a diagnosis. At that point a global evaluation function is run. It tries to determine if there is a Happy frame which can account for all of the abnormal findings. (A list of the findings is created by the input mechanism.) A finding is accounted for

by a happy frame if it is a slot in that frame or if it is accounted for by a slot in that frame (watch the recursive definition!) Here is where comp1ication pointers become useful. If A is a complication of B and if C is a finding of A then if B is happy and A is happy, C is accounted for by B. (If we think a person has AGN and also Nephrotic Syndrome we can account for his lipiduria by the hypothesis of AGN because Nephrotic Syndrome is a possible complication of AGN.) Clearly this theory is not yet very well developed but I believe that it is on the right path.

III A Suggested Protocol for Research

Because of the vague and tenuous nature of the theory, I would caution against premature programming except for experimental purposes to check out ideas. I feel that before any serious attempt is made to code a program which we consider a basis for any future system of diagnosis and management, we should firm up this theory and fill it out with more detail and examples -- or throw it out and replace it.

One high priority item should be the development of a language, perhaps an elaboration of the graphical system of which an example has been shown, whereby medical doctors who are totally unaware of programming techniques, can describe medical concepts, such as glomerulitis, and their relationships to other concepts in a precise and mechanistic form. I feel that it would then be of value for us to completely map out some small section of medical knowledge, say the glomerulidities via this mechanism. Probably many new issues would be raised just by this exercise. We must also create new linguistic

structures to enable the representation of the local evaluation
functions, so that they can be represented. (Perhaps a start has been
made by introducing several different kinds of evidence pointers.)

After we have constructed the map of the glomerulidities we should
hand-simulate it on several (made up?) CPC's to get a feel for what kinds
of processing goes on. Only then should we build a program to absorb the
graphical representation and use it to manipulate facts.