Helping with a Humanly Impossible Task: Integrating Knowledge Based Systems into Clinical Care

A.L. Rector
Medical Informatics Group
Department of Computer Science
University of Manchester
Manchester M13 9PL
Telephone: +44-61-275-6188/6133 Fax: +44-61-275-6236
JANET: rector@uk.ac.man.cs.ux

Citation Information:
Originally appeared as:
Helping with a humanly impossible task: Integrating knowledge based systems into clinical care. Published in J Hannu and L Seppo (eds), Proceedings of the Second Scandinavian Conference on Artificial Intelligence SCAI-89. Amsterdam: IOS Press, pp 453-466

A shortened and revised version appeared as:

Introduction
Medical Systems - why not in use.

Medicine has proved a fruitful field for developing knowledge based systems. The existing knowledge sources are relatively systematic; there are well defined subproblems of a tractable size; and it has been possible to recruit interested experts.

Paradoxically, the medical environment has a number of characteristics which make it difficult to introduce knowledge based systems into routine use. Doctors work in a complex social environment under severe time pressure. A large number of people must agree to use a system before it can be successful. The neatly confined problems most tractable to expert systems often have limited relevance to doctors' decision making in practical situations. Furthermore, doctors already have a well developed system of referrals and sharing of expertise, so that there is no glaring gap waiting to be filled by an expert system. To these problems must be added the financial pressures which have limited the power of the hardware and sophistication of the software and concerns about validation and safety.

As a result, the effort invested in developing medical applications of artificial intelligence has not so far had a major impact on routine clinical care. Attempts to produce systems for other professional users - e.g. architects, lawyers, accountants, executives - have had somewhat similar if less dramatic experiences, despite the current excitement about ‘Executive Information Systems’

The apparent reluctance on the part of doctors to accept ‘expert systems’ may also be an example of a related phenomenon. Our experience of introducing computers in hospitals is that word processors were taken up first by academic consultants and by staff in departments which were short of secretaries. Most secretaries accepted them only reluctantly, at least at first. Similarly, we found that spread sheets were taken up instantly by departments such as catering, estates management, and the various laboratories, but that it took much longer before the finance department made extensive use of them.

We would put forward the hypothesis, that it is much more difficult to introduce a system which impinges on users’ primary skills and expertise than systems which
assist with peripheral, probably irksome, tasks. For the academic, retyping a page is a major chore; for the catering manager, refiguring a budget is a major chore. For the trained secretary or accounting clerk, respectively, these are but a few minutes routine work exercising a skill in which they take pride.

If this hypothesis is correct, then it is hardly surprising that doctors have been slow to take up systems aimed primarily at diagnosis - the task most central to their expertise and one which they find interesting if not always easy.

Payoffs for everyone. /Involve users

Is there then a role for knowledge based systems in medicine? We are trying to step back and examine what practising British general practitioners feel are their major needs and to involve groups of them intimately in reassessing how those needs might be met with the technology now becoming available.

The design which is emerging from our work emphasizes cooperative, mixed initiative problem solving. It embodies a view of 'Artificial Intelligence' as 'Building artifacts compatible with intelligent users'. The design uses many techniques from AI, but owes as much to the literature on the computer support for cooperative work (e.g. [4]), data base management systems, and human computer interaction. While based specifically on the problems faced by British general practitioners, we believe many of the features are more generally applicable. The process is still in its early stages. The paper that follows traces our motivations and some of the basic principles which are emerging.

Computer Support for Professional Activity

Common Issues

We distinguish between systems which augment skilled performance and systems which substitute for some skill or knowledge. Classic expert system paradigms such as MYCIN or INTERNIST substitute for an expert. Perry Miller's critiquing systems ATTENDING, ONCOCIN, the recent adaptation of INTERNIST to become QMR and the Oxford System of Medicine are steps towards systems which might augment experts' abilities rather than replace them.

Substitution systems assume that their users are ignorant, or at most novices, in the field. Augmentation systems assume that their users are 'broad experts' who are skilled in the field and exercise ultimate judgement, although they may make slips or lack particular items of knowledge.

Augmentation systems are used whereas substitution systems are consulted. While this distinction is not absolute, the goal of most augmentation systems is to become part of the regular tools of the trade, just as the other tools which are ‘readily to hand’ in the users’ environment. For these reasons we prefer to think of our work as designing a ‘human computer environment’ rather than a ‘human computer interface’.

We believe that it is not just a matter of making the systems easy to use but of making them useful to their users. And we take as our starting point that only the users can judge whether or not a system is useful. (We contrast the ‘usefulness’ of the system which can only be judged by the user with the ‘benefits’ of the system which may accrue to users, clients, the organization, or society at large.)

This is not to downgrade the importance of a good user interface. To the contrary, one of the major contributions of Ehn at the Sweedish Institute for Working Life and of Winograd and Flores is to point out the insight, which goes back to Heidegger, that we only notice a tool when there is a ‘breakdown’ in its use. By this criterion, most users would find most existing systems ‘broken down’ most of the time.

Executives, lawyers, doctors, - problems differ but issues similar
Links to work on CSCW. Collaborative models. Sweedish Inst for Working Life
What kind of system...  
Distinction between augmentation of expertise and substitution of expertise/ and between autonomous systems and directed/monitored systems.

Speculate on the role of training in ESs.  
User Interface and User Environment  
cite Fox and Rector amongst other things

Special problems of doctors

Engelbart coined the term ‘knowledge workers’ to cover a wide variety of professionals. However, he, along with others in the same tradition such as Winograd, have considered ‘knowledge workers’ primarily in the context of business management.  

* Doctors’ primary attention is on the patient rather than the system. An engineer using a CAD/CAM system, or perhaps an executive using an executive information system, will sit down and use the system intensively for a period ranging from a few minutes to a few hours. During that time their attention will be primarily on the system. Doctors’ primary attention should be focussed on the patient. The system cannot be ‘to hand’ unless it is present during the consultation, but the patient must remain the focus of attention.

* Doctors work in brief consultations under severe time pressure. In most medical fields average consultation with a patient lasts only five to fifteen minutes; in British general practice, the average consultation is only seven minutes. There is little long term planning. The focus of each problem solving episode is to find an immediate solution to the current problem and ‘get the patient out the door’.

* The variety of problems with which any general practitioner is presented in the course of a week is enormous and no one condition occurs sufficiently frequently to justify a specific system to deal with it. For example, even diabetes, which affects between one and three per cent of the population, presents a typical British general practitioner with a troublesome decision problem only once every few days. Nonroutine cases of less frequent but still common diseases, such as epilepsy, may occur on the order of a few times per month.

* Doctors already work under an information overload. There is every danger that any information system will simply aggravate the overload.

* General practitioners work in comparative isolation from their colleagues. Much of the ‘training’ in how to use systems in other fields occurs through informal contacts between users. The structure of medical practice makes such informal contacts less frequent.

* There is a wide diversity in users’ clinical skills. Medical practice includes many transient doctors - housemen and registrars in training, locums filling in, junior partners who join and then leave, etc. Each doctor has an idiosyncratic range of clinical skills, with hard won expertise in some areas and serious lacunae in others.

* A wide range of experience with the system is to be expected. If a system really is successful at becoming a routine part of practice, many doctors will be using it four to six hours per day, four and a half days per week. Like any other experienced user they will require short cuts and every second’s delay will be a significant irritant. On the other hand, the system must be accessible to locums, trainees, and new partners who meet it virtually for the first time.

* Although most encounters with patients occur in the consulting room or at the bedside, there are many variations. Home visits, ‘bed borrowing’ in hospitals, casualty wards, and specialist laboratories all have special needs. Experience
with conventional information technology in medicine gives many warnings that failure to satisfy critical variations can make otherwise desirable systems unworkable in practice.

* Psychological and social issues are an integral part of medical practice, and much of the information needed to deal with them is ‘soft’, poorly characterized, and therefore difficult to evaluate.

In addition there are three difficulties which tend to make inappropriate the type of diagnostic expert system which has been most frequently developed:

* Most clinical effort is not spent in diagnosing patients de novo. When problems in decision making occur, they involve management at least as often as diagnosis, and they more frequently involve some specific piece of information than the global evaluation of the patient. In fact, in many cases the doctor will already know what the most likely answer is, and his or her question will be about one of the less likely but worrying possibilities or about a general feeling that “something is not quite right”.

* There are few obvious gaps in the available network of sources of expertise waiting to be filled by ‘expert systems’. In fact, patterns of referral and consultation are fundamental parts of the social and financial framework of the medical profession. The existing structures have been developed over the one hundred and fifty years of the development of modern medical practice to ensure that there is always a suitable colleague or consultant from whom to seek advice. There is a natural resistance to anything which appears to threaten these relationships.

* Although most consultations follow a stereotyped script, few systems take full advantage of this fact to make the information or actions likely to be needed next readily to hand.

Human and Medical Fallibility
Sources of Error: Slips, Muddle, Mistakes, and lack of information
Reason/Rasmussen
Distraction
Overloading

All of these features would seem to make medicine a singularly poor candidate for computer assistance, and they help to explain doctors' resistance to ‘expert systems’. On the other hand, from a broader perspective there is ample evidence that the practice of care could be improved. Furthermore studies of human error in other fields suggest that medical practice is almost ideally designed to promote errors.

Doctors work in an environment where much of what they do is routine but in which they are never fully in control. Patients constantly change the agenda by bringing up additional problems. In addition to the routine task, there is usually some other detail to be dealt with. Given this situation it is unsurprising that studies of medical audit, (e.g.) show consistently that doctors frequently omit simple checks and make simple errors in prescribing despite adequate knowledge. In fact, the classic study of this phenomenon by McDonald showed no association between the level of training or perceived skill of the doctors and the frequency with which they made routine slips which could easily be audited with the help of a relatively simple computer system. The problem is not simply one of knowledge skill or attention, it is the result of intrinsic features of clinical practice and how they interact with the limitations of human cognitive processes.

However, it is also true that the range of knowledge required to deal successfully with all problems in general medicine is too great for any one person to encompass. Not
only are slips inevitable, but all doctors will have areas in which their knowledge is 
faulty so that mistakes are likely.

**Principles of Design for a Medical Decision Support System**

**Structural Issues**
- Defining appropriate roles
- Redefinition of professional roles
- Pride in craftsmanship
- Identifying the real tasks
- The normal trajectory of events
- Inappropriate Chunking
- The importance of variances

**Basic Principles**

How might these problems be overcome? Our approach is based on:

* A broad vision of a ‘clinical information environment’ based around the medical 
  record, which we hope will become doctors’ primary means of gathering and 
  retrieving information and of communicating with the rest of the health service.

* The determination that a sufficient number of doctors will be involved in the 
  design of the system that it will reflect the real needs of a significant segment of 
  the medical profession.

The involvement of doctors in the design process is critical. It dictates an 
‘evolutionary’ approach through rapid prototyping and experimentation, since neither 
doctors nor designers can predict with any accuracy how new ideas will work in 
practice, and doctors can only explain their needs in response to examples of the 
possibilities.

**Key Maxims**

Within this framework, we have been conducting a series of workshops with general 
practitioners and observing them during consultations, as well as examining a range 
of systems aimed at doctors in various specialties. Out of this work and the more 
general considerations outlined above have come a series of maxims to guide the 
design:

* Make the system useful to doctors in day to day patient care.

* Do not overload the doctor. One of our central efforts is to find means of 
  summarizing and ‘chunking’ information so that it can be assimilated quickly 
  within the limits of human short term memory. The top level of information 
  should cover only the detail normally needed. Further information can be sought 
  by expanding individual items. Academic completeness can be the enemy of 
  usability.

* Keep the number of interactions with the system to a minimum. Doctors’ 
  attention should be on their patients; having to manipulate the system is a 
  distraction. Reducing the number of interactions conflicts with conventional 
  wisdom that the amount of information on the screen at any one time should be 
  kept to a minimum. Finding the right balance between the amount of information 
  on the screen at one time and the number of interactions required to access or 
  enter a particular piece of information is one of the functions of the experimental 
  prototypes.

* Keep the extra memory load on the doctor to operate the system to a minimum. 
  Use direct manipulation metaphors and make the navigation through the system 
  visually obvious.
* Make the system flexible and tailorable. Different doctors and groups of doctors will have different requirements. Medicine is not a precise science, and individual variations range from substantive issues such as preferences for different drug regimens and disagreements about the use of diagnostic tests to surface level issues such as whether a particular doctor prefers icons or text in a given context and whether or not he or she is willing to use the keyboard.

* Make the system quick and efficient to use for the common situations which constitute the vast bulk of clinical care. Do not allow the system to become overburdened with things which things needed only rarely.

* Examine the normal ‘trajectory’ of the consultation and try to make the system fit it smoothly. Insofar as possible, the ‘next step’ should always be natural and immediately to hand.

* Let the doctor do anything at any time and pursue as many threads as needed. The course of a medical consultation cannot be predicted in advance. Doctors must be free to initiate a new line of action from any point in the consultation.

* Never require that the doctor do anything. All data fields and all actions should be optional.

* Tailor the system to the user. No one interface will fit everyone all of the time.

* Always look for the disadvantages of using the system and the potential harm it might do as well as the potential good.

An intelligent medical record - information readily to hand.

Where should we start? What does having information readily to hand mean to the practising doctor? The basic information used by doctors is about the patient, and so our design starts with the basic source of patient information - the medical record. How to reform medical records has been an issue in medical circles for the past 25 years, but attempts to do so have met with only limited success. A patient with a complex history may have a paper record several inches thick, minimally indexed, much of it handwritten, with no reliable summary. The task of extracting a complete coherent history from a set of clinical notes can take a doctor over an hour, and so is rarely done. Errors which occur because doctors fail to find important information are common. Attempts to provide summaries manually such as the Problem Oriented Medical Record have usually floundered because of the extra effort to make duplicate entries in the summary and the body of the record. Furthermore the summaries themselves rapidly become so cluttered that they are no longer of any use.

Most existing computer based medical records do little more than create an electronic equivalent of the paper record, a simple chronological record of what has occurred. Since the record of a patient with a complex medical history may contain over 10,000 ‘observations’, an uncatalogued or poorly indexed list is hardly helpful. Our experience in going over general practice records of such patients - a process in which one of the doctors in the practice goes through the notes and summarizes them for us verbally while we take notes - is that there almost always comes a point of acute embarrassment where it is painfully obvious, even to the medically untrained, that the doctors were in a muddle and had lost track of the situation. (Usually, and some later point, the patient is referred or one doctor takes it upon him or herself to ‘sort out’ the patient.)

The task which we have set ourselves is to provide a self-organizing and well summarized record which can be viewed in a variety of ways. Much of the ‘intelligence’ is devoted to making the data entry in the record as easy and smooth as possible. The goal is to replace as much as possible of the paper record so that they system becomes truly a part of routine practice. We expect most of the ‘decision support’ in the system to come from the effective presentation of information and from
implicit prompting by data entry formats. Our metaphor is of improving the ‘light’ on the information so that the doctors can see better and so perform better.

**Summarization and Chunking: An expert system to answer the question “What next?”**

Fundamental to the design is the need to answer three questions:

1) What information is the doctor likely to need to record about the current situation?

2) What one ‘screenful’ of information does the doctor need currently - and how can it be presented so as to be assimilated easily?

3) Is anything which the doctor is proposing or omitting sufficiently worrying to merit warning the doctor?

Successful answers to questions one and two both involve establishing the appropriate ‘chunks’ for medical dialogue. One of the chronic complaints about many existing systems is that they demand too much detail. Doctors usually communicate in a shorthand of syndromes or ‘super syndromes’, for example ‘pneumonia and heart failures’ or ‘disabling migraine’. In most situations this is all the information which is needed; it may even be all that it is appropriate to record. Clearly a distinction must be made between what can be taken to be true by default and what is actually recorded, and provision must be made to revise the current beliefs as assumptions are found to be false. But this is just the general problem of default reasoning and belief revision, the subject of an entire subfield of AI, placed in the context of a specific application domain.

To limit keyboard entry, most entry is done from constrained selections - pointing, menus, and direct manipulation. For this to be successful, the user must have the options presented which are most likely to be needed. To achieve this requires a visual language and something approaching semantic and pragmatic grammars of clinical practice. Fortunately, much of medical knowledge is inherently visual, and we are finding the basic association of a functional system with an anatomical region a surprisingly powerful basis for the visual language.

The distinction between the semantic and pragmatic components of the underlying ‘grammar’ is critical to the system. The semantic component answers the question “What can be said?”. The pragmatic component answers the questions “What is likely to be said?” and “What needs to be said?”. Data entry is organized in layers, such that the doctor can always ask to enter more detail than is currently provided. The top layers follow the pragmatic component of the knowledge base. If the pragmatic component were totally successful and if the world were completely predictable, this would be all that was needed. However, inevitably from time to time the doctor will need to get beneath the level of what is expected to the level of what is possible. How often this will occur in practice remains to be seen.

Our early work on summarizing information for display indicated that a few simple heuristics were surprisingly successful in summarizing clinical notes within limited fields. The heuristics were based on the established taxonomies of clinical problems supplemented by a modest amount of explicit knowledge about the relations between conditions. We remain hopeful that this approach will prove successful in a larger study, but this aspect of the work is still in its earliest stages.

**Architecture of the System**

**Structured Meta Knowledge**

The architecture of the system is based on representing both the knowledge base and the medical record as a semantic network with inheritance using the methods
developed in our prior work on ‘Structured Meta Knowledge’\textsuperscript{24, 25, 26}. We view the medical record as a set of ‘descriptions’ of patients, or rather of ‘patients as seen at a particular time and place’\textsuperscript{27}. From the point of view of the standard expert systems architecture, the medical record is seen as the ‘dynamic data base’ and the ‘categories’ (classes) in the semantic network as the ‘knowledge base’. The observations in the medical record are treated as ‘occurrences’ (instances) of the categories in the semantic network.

There are two key features of the structured meta knowledge approach:

* The arcs in the network, as well as the nodes, are first class objects and can be described. In particular the explanation or evidence supporting an arc can be ascribed to that arc. This allows a) layering of the knowledge base in which surface statements are explained by deeper statements; and b) reason maintenance based on explicit structures in the network rather than an additional primitives implicit in the inference mechanisms. Making the reason maintenance explicit makes it possible to consider which parts of the reason maintenance network should persist or be communicated, e.g. how much of the reason maintenance structure should be transmitted when exchanging medical record electronically between different systems, how much should be stored permanently with the record for use at subsequent consultations. Since the knowledge base will inevitably change over time and different centres have different versions of the knowledge base, these questions are nontrivial.

* It is ‘generative’; \textit{i.e.} it allows new concepts to be constructed out of the existing concepts, for example from ‘pneumococcus’ and ‘pneumonia’ to construct ‘pneumonia caused by pneumococcus’. One of the reason for the enormous size (up to 250,000 terms) of standard medical classification and coding systems is that they have limited, if any, ability to construct concepts and therefore require an vast array of ‘primitives’.

Object Orientation and Layering
Programming as reconciling the mismatch between the magic number $\pm 7$ and the creation of artifacts with thousands-millions of interacting parts.

The relations within the knowledge base form one axis of the system. The relationship between the categories in the knowledge base and the occurrences in the medical record form a second axis. The third axis of the system is formed by the layers of the application and user interface:

- Displays
- Presentations
- Perspectives
- Things

The bottom layer, the things are the objects in the semantic network itself. The thing layer is responsible for maintaining basic integrity constraints and has a straightforward translation into standard data base structures.

The perspective layer provides a) filters and transformations so that the things can be viewed in different ways; and b) interpretation of the commands from the presentation layer into operations on the thing layer. Most of the ‘intelligence’ of the system lies in the perspective layer. Different perspectives perform summarizations, decide on appropriate chunking and levels of abstraction, and restructure the information into problem-oriented, chronological, or hazard-warning views.

The presentation layer manages a) the ‘topological’ information concerning presentations - i.e. information which is relatively independent of the details of the screen itself such as the tabbing and indentation structure of a piece of formatted text; and b) the conversion of the ‘tokens’ from the display layer into ‘commands’ to the perspective layer.
The **display** layer manages the actual mechanics of the screen display and the conversion of ‘lexical’ input from the user into ‘tokens’ to be dealt with by the **presentation** layer. In the Smalltalk80 implementation, the **display** layer is a view-controller pair.

This multi-layered model serves two functions and has features in common with other layered models discussed widely in the HCI community (e.g. [28](#)). It allows the complex interface to be factored into pieces which it is practical to implement and maintain. It also allows multiple views of the same information to be maintained simultaneously and consistently. There can be any number of perspectives on the same **thing** open simultaneously and any number of **presentations** of the same **perspective**. (In theory there can also be more than one **display** of a **presentation**, but we have not been able to find any use for such an arrangement.) All changes are made ultimately in the **thing** layer and the effects are propagated back to the multiple presentations via the dependency maintenance system. The display of any individual patients record consists of a number of windows each of which **displays** a **presentation** of an alternative **perspective** on that patient.

**Conclusion**

Professionals have a wide range of skills only some of which are easily captured by the formal descriptions of their disciplines. They carry out a wide range of tasks, only some of which fit formal or academic stereotypes. While not sufficient, the formally described tasks are a necessary part of performing effectively, and limitations on human cognitive processes make errors inevitable. However, professionals are not novices; they are broad experts, and systems which treat them as novices are likely to be rejected.

From our current work, we can see at least three different ways of augmenting the abilities of one group of professionals - doctors in general practice - to carry out their tasks:

* Performing tasks such as retrieving, sorting, transforming and displaying information which ‘throw light’ onto the problem in question and provide cues which would otherwise be missing from the environment.

* Monitoring and anticipating the situation according to the formalized rules of the discipline to anticipate or correct errors.

* Performing particular tedious subtasks which are outside the doctors’ primary expertise.

To achieve any of the three, requires systems which the doctors themselves find **useful** in their daily tasks and which are matched to their usual practices and concepts. The work is still in its early stages. Ultimately we hope that the systems might become ‘intelligent assistants’ which make ready to hand a wide range of services, some of which will resemble knowledge based systems, others of which will be more mundane.

---

**Acknowledgements**

This paper is based on the collaborative effort of the Manchester Medical Informatics Group. The research is supported in part by the Medical Research Council and the Department of Health.


21 Weed LL (1969). *Medical Records, Medical Education and Patient Care.* The Press of Case Western Reserve University, Cleveland, Ohio.


27 Rector AL and Kay s (1989). Descriptions???
