PHD

The machinery of medicine

Hartland, Joanne

Award date:
1993

Awarding institution:
University of Bath

THE MACHINERY OF MEDICINE:

An Analysis of Algorithmic Approaches to Medical Knowledge and Practice

Submitted by **Joanne Hartland**
for the degree of PhD, 1993
University of Bath

J. Hartland.

This work is dedicated to the memory of

VIVIEN NEIL

THE MACHINERY OF MEDICINE:

An Analysis of Algorithmic Approaches to Medical Knowledge and Practice

TABLE OF CONTENTS

ABSTRACT

To what extent can machines and computers take over from people? Will they render us all redundant? This research answers these questions, using 'the machinery of medicine' as a vehicle for the investigations.

Designers of smart machines utilise an algorithmic model of knowledge. This raises questions for sociologists who stress the enculturational model of knowledge and the importance of the tacit component of skilled performance. How effective is the algorithmic approach to medical knowledge and practice?

I have adopted an 'insider' approach in order to highlight the practical issues of the usefulness of medical machines, their 'competence' in the eyes of the staff that use them, and the contribution of these staff to the efficient functioning of the machines.

The fieldwork shows that the machinery of medicine relies on behaviour-specific action, and functions best in 'microworlds'. Problems arise when machines are assigned to wider world tasks that have been inappropriately reduced to a microworld format. I conclude that human digitization and repair allows microworld machines to fit into the social world of situated action. I explain where and how medical machines offer most benefit, consider whether they should be used by inexperienced staff, and I suggest where designers should concentrate their efforts. The research shows that the notion of a 'post-physician era' is misplaced: Machines cannot approach the broad aspects of medical practice, and the human role remains essential to the appropriate use of medical machines.

ACKNOWLEDGEMENTS

Practically, I am indebted to the medical staff and the patients who provided assistance far beyond the call of duty: The Cardiology staff at the University Hospital of Wales, at Llandough hospital in Cardiff and Llwynypia hospital in the Rhondda; the medical staff at the Research Institute for the Care of the Elderly at St. Martin's hospital, Bath; the Gastroenterology firm at Llandough hospital; the medical team of February 1992 at the Whittington hospital, London and the Medical Informatics staff at the LDS hospital, Salt Lake City. I also thank Jo Schol for her patience and practical help over the last few weeks.

'Academically', Harry Collins's influence has been especially important. His patience, wry criticism, and above all his enthusiasm, have been inspiring. I greatly appreciate his extremely skilful supervision and his thoughtful advice on all matters. It has been a pleasure working for him. I am grateful to other colleagues - Marc Berg, Barrie Lipscombe, Mike Kelleher, Paul Atkinson, and the 'critical mass' of post graduate students - Gill Hatt, Steve Travis, Warren Evans and Robert Evans. They have read and dismantled my early efforts, endured my anguish and shared my joy at minor successes.

On a personal level, friends and family have provided all kinds of assistance. They have sympathised, given free advice and liberal quantities of wine, made me laugh, loaned me money and even listened to Bob Dylan with me. Gavin Nobes, Judy Wright, Gill Tipping, Judith Turbyne, Mark Ellison, Hazel Wallis and my family - Gwyn, Jean and Ruth Hartland.

Finally, Tony Wilton has given practical, academic, financial and personal support. Thank you for your unorthodox methods of encouragement and stimulation, which have been extremely effective, and which I will not forget.

LIST OF FIGURES

GENERAL INTRODUCTION

"In gadget-mad Japan, they have buildings that think ... The NEC building [in Tokyo] is apparently the most intelligent in the world" (1)

The idea of a 'thinking building' shows how far images of intelligent machines and smart computers intrude into our lives. In the developed societies of the late twentieth century, machines are everywhere. At one end of the spectrum they are employed to do regimented or straightforward tasks - on factory assembly lines and at supermarket check-outs. At the other end of the spectrum machines are used for more complex operations - they manoeuvre aeroplanes by autopilot and facilitate high speed, large scale calculations and communications. In this mechanised environment the image that the machines present is one of relentless motion and tireless, near-perfect operation.

Winston and Prendergast (1984) tell us that "The primary goal of AI [Artificial Intelligence] is to make machines smarter." (p 1). The ever-improved versions of the kinds of machines mentioned above suggest that the designers are achieving a degree of success. The age of the computer is with us and the influence of the machines continues to expand. There are hundreds of computers that seem to do what were once considered 'human' tasks. Researchers strive to make machines that relieve us of more and more of our time-consuming tasks. Machines have influenced all sectors of society from leisure and pleasure to finance, travel and education (see Winograd and Flores, 1986, chapter 1). Medicine and health care has not escaped - as Koenig (1988) points out, "The landscape of modern health care is filled with machines." (p. 465)

The practical expansion in the influence and scope of computers has been accompanied by heated debates about the nature of computer capability, the limit of this capability and the future of AI. Philosophers, psychologists, sociologists, natural scientists and engineers have all contributed to these debates. In addition, the production of allegedly expert computers that threaten to encroach upon professional boundaries has aroused interest amongst human experts from a diverse range of occupations.

Analyses of the potential of AI have been conducted largely through theoretical debate. Supporters of the ideals of the AI community, such as Feigenbaum and McCorduck (1984), and Maxmen (1976, 1987) have made bold predictions about what computers will achieve. Sceptics have countered these claims, pointing out what computers can't do (see Dreyfus 1972, 1992) and what computers ought not to do (Weizenbaum 1976). Some writers have attempted to explain computer successes to date, analysing what computers can do and why they can do it (Collins 1990). Others have chosen either to take a pragmatic stance or to adopt a mobile position, changing their views as the area has matured (eg. Schwartz 1970; Schwartz Patel and Szolovits 1987; Szolovits 1982).

The debate is largely inconclusive, but the rhetoric serves to underline the differences that exist between proponents and opponents of AI. Much energy has also been invested in debating questions of intelligence - whether machines can be 'intelligent' and whether we should grant them 'intelligence' when they perform jobs that humans once did. But little attention has been given to practical investigations of machines that are doing what were previously regarded as human jobs. A practical investigation offers a means of circumventing the issue of whether the machine is intelligent, by asking instead, how well does the machine do what it is designed to do, and how does it do this? Whether or not the machine is 'intelligent' is irrelevant if it performs in a satisfactory manner.

This thesis presents such a practical analysis using medical computers as a vehicle for the investigations. There are two reasons behind the choice of medical machines: First, within the medical world there are both simple devices and more complex computers. Second, a variety of approaches to mechanical decision making that AI researchers have devised are represented in medical machines. Rule based systems as well as computers that combine rules with specific knowledge about the field in question have been introduced. So the field offers opportunities for analysing several different types of machines that are believed to be 'successful' - thus Forsyth says, "The most significant thing about expert systems is that they are highly successful: already there are systems that can out-perform skilled humans at medical diagnosis." (1984 p. 9) These knowledge based systems, or 'expert systems' are currently the most popular type of machine in practical medical settings (2). Medicine is, then, a fertile field for looking at the practical issues surrounding machine use (3). The field studies in this research show how machines are used, their usefulness, the kind of tasks they are allowed to undertake and the effect their introduction has on the human workers involved. The discussions and arguments arising from these case studies will be presented in a manner that will, I hope, allow an informed reappraisal of current positions and debates.

The major objectives of this study are:

(1) To assess current theories about the extent to which machines can be used to do human tasks.

(2) To investigate the concept of machines performing human tasks using case studies from the medical world.

(3) To offer a critical reappraisal of current theories and present new information.

Before presenting the case studies, a review of the literature in the area will be presented. The structure of the thesis is as follows:

Section 1 - chapters 1-4 - will present the background to the field, dealing with theoretical, historical and practical issues raised in the literature. This section sets the scene for the empirical research.

Chapter 1 provides an outline of Sociology of Scientific Knowledge (SSK). The evolution of SSK is described, the SSK conception of the nature of scientific knowledge is explained and the thrust of new ideas within SSK outlined. How SSK relates to the area of AI, and the bearing that the new radical strand within SSK has on AI is then discussed.

In chapter 2 the growth of AI, and how the direction of research within the field has evolved is described. AI in medicine is singled out as the case under scrutiny and the available literature reviewed. This chapter illustrates how the problems facing AI researchers generally are reflected in attempts to produce machines to do human tasks in medicine.

Chapter 3 completes the literature review with attention to another practical issue raised in the literature: The effects of technological advances and the implications for the human workforce have always been an emotive issue. For all workers, from the Luddites to the Toyota workers of the 1990s, the introduction of technology has posed a threat to autonomy, skill, status and to employment. In this chapter Braverman's deskilling thesis is reviewed and alternative interpretations assessed: Other writers suggest that the extent of deskilling is far more limited than Braverman's pessimistic account forecasts. At this point, the link between the notion of deskilling and the growth of computerisation is made, and various theories concerning what computers and smart machines can and cannot do are set out.

In chapter 4 the specific points raised by the three sections of the literature review are summarised and the questions that the case studies need to address are set out. These questions replace the 'hypothesis' that is often formulated for testing at the early stages of research. Instead the questions will address the difference between the SSK conception of knowledge as social and the alternative ideas implicit in the AI paradigm, through a practical programme of investigations. What happens in practice when computers are employed to do human jobs in health care settings is the major issue. The case study material will allow a practical reappraisal of current theories, and the chance to contribute to the debate from a different angle. At the end of chapter 4 the methodology used for this research is outlined. Why this method was adopted, how it was implemented, the problems of qualitative research of this nature and the fieldwork practicalities are explained.

Section 2 - chapters 5-9 - reports on the practical elements of the thesis. Each of these chapter is designed to provide the background to a specific case study, a description of the particular machine involved, an outline of the findings and an analysis of these findings in relation to the questions raised in chapter 4.

Chapter 5 gives details of a computer designed to interview and diagnose patients with dyspepsia (GLADYS). In chapter 6 a computer that diagnoses jaundiced patients on the basis of clinical information (SOLUBILE) is described. Chapter 7 gives details of the study of the automatic blood pressure measuring machine. In chapter 8 the investigation into the interpretive electrocardiogram machine is presented and chapter 9 is concerned with specific applications of the 'Dynamic Hospital Information System' (HELP) in use at the 'Latter Day Saints Hospital' (LDS) in Salt Lake City.

Section 3 - chapter 10 - is the conclusion of the thesis. Here I summarise the findings of the five case studies and using this information I make suggestions about the extent to which computers contribute to medical practice, where computers will be most useful, the role of the staff and their contribution to the new working practices. The way that this practical investigation has contributed to the field is explained. Conclusions about the usefulness of current theories are presented; a new practical way of looking at the issues will have been fully explored and the implications of the findings for all the staff involved is explained. Suggestions for the improved use and application of machines in medicine are set out and the direction that further studies should take in order to advance the discussion is outlined.

FOOTNOTES TO THE INTRODUCTION

1. Margolis (1993).

2. Neural networks, hailed as the next step after expert systems, remain at the research front. Few, if any, examples of neural nets being used in practical settings can be found. There are none deployed in medical settings at the time of writing.

3. Aside from the variety and apparent availability of machines to study, this area is a fertile field in which to pursue this investigation because of the nature of my expertise in the area. This is explained in more detail in the methodology section of chapter 4.

SECTION 1

CHAPTER 1

THE SOCIOLOGY OF SCIENTIFIC KNOWLEDGE MEETS ARTIFICIAL INTELLIGENCE

CHAPTER 1 - THE SOCIOLOGY OF SCIENTIFIC KNOWLEDGE MEETS ARTIFICIAL INTELLIGENCE

1. INTRODUCTION

The past quarter century has seen the establishment of the Sociology of Scientific Knowledge (SSK) as a distinct discipline. Areas as diverse as the history of science, philosophy, the sociology of knowledge and the sociology of science have contributed to the development of SSK. In this chapter I describe the evolution of SSK and the various methodological positions that have been put forward. Then the link between Artificial Intelligence (AI) and SSK is made clear. The expanding influence of Artificial Intelligence and smart machines offers a fundamental challenge to SSK and in view of this, I explain why a 'sociology of machines' is justified.

1.1 The Evolution of Sociology of Scientific Knowledge

Historically science has been held in high esteem and generally regarded as a reflection of objective truth and fact. This 'rationalist' or received view positioned science beyond the remit of sociologists of knowledge, who were traditionally limited to the analysis of false or irrational beliefs. As Woolgar (1985) pointed out, "While science was generally regarded as exotic and esoteric, it was neither necessary nor desirable for the sociologist to penetrate the content of science." (p. 559). Woolgar (1988) suggests that the rational view of science is inextricably linked to what he terms 'essentialism'. Essentialist views are based on the idea of science as a tangible object - exhibiting features that are characteristic of an 'actual' existing science. Woolgar presents nominalism as an alternative to essentialism. Nominalism is based on the notion that the search for a definition of science is futile because science is variable - philosophically, culturally and historically. It is continually changing, being re-negotiated and re-classified by the involved scientists. The characteristics of scientific knowledge are thus more dependent on the practices of the *people* involved, than on the characteristics of science itself. As interest in the practices and social relations of scientists grew and it became accepted that these social factors influence what is counted as 'scientific' at any time, the sociology of science became established as a means of investigating the social relationships (1). Writers such as Merton were amongst the first to approach this area (2).

The Sociology of Scientific Knowledge grew from the sociology of science, but its genesis required another major step forward: Rather than concentrating on the social relations between scientists, 'SSKers' look - from a sociological perspective - at the nature of scientific knowledge and the way it is produced, disseminated, maintained and passed on within the scientific community. As long ago as 1967, Medawar urged those interested in understanding science to look at what scientists *do* in their research (p. 151), and take note of the social dynamics at work. To study these influences in action, close attention to scientific practice itself is required.

Kuhn (1962) was one of the first to present ideas which suggested that science is a 'social' practice. Kuhn argues that knowledge production and transfer depend on two factors - paradigms and exemplars. 'Paradigms' mediate what is regarded as 'fact'. In order to become part of the science community, allegiance to the dominant paradigm is necessary. This is achieved by interaction with and immersion into the social/scientfic group. In this way, the frameworks referred to inside the group are passed on to newcomers. 'Exemplars' are previously experienced concrete examples which act as judges over the structure of our world view. Kuhn maintains that there is no simple objective knowledge of the world, but that accumulation and transfer of knowledge depends upon paradigms and exemplars. Kuhn's work, along with the earlier work of Polanyi (eg, Grene ed, 1969) illustrates the craftsmanlike character of scientific work and emphasises the social influences shaping the outcome of scientific experiment.

This emphasis on the role of experience and reference to shared ideas, or frameworks of a paradigm, bears a resemblance to the Wittgensteinian concept of a 'form of life' (1958) (3). Winch (1958) also utilised the concept of a form of life. He maintained that we come to know what we know and we act as we do because we share a common form of life. Within a form of life, knowing and doing cannot be separated. This applies to the science form of life, as much as to any other form of life.

From these ideas, it became clear that science, like all other bodies of knowledge, is amenable to sociological analysis and that "the historical course of development in any field of science has a significant social component." (Ziman 1984, p. 103) By highlighting the social nature of accomplishments in mathematics, physics and the like, 'SSKers' hoped to show that the epistemological supremacy of science is mythical. The case studies chosen in the early days of SSK were typically hard cases of science such as mathematics (Bloor 1973) and physics (Collins 1975). The sociologists involved set out to demonstrate that the apparent objectivity of all science is a function of social processes (4). According to Barnes (1974), the study of scientific knowledge also assists the general quest for a sociological understanding of the nature of knowledge as a whole (p. viii).

1.2 The Relativist view replaces The Rationalist view of Science

Studying the production of scientific knowledge sociologically is at odds with the rationalist or received view of science, which treats the social aspects of science production in a much less thoroughgoing way. SSK demands that the received view of science be abandoned in favour of a more relativist stance. The distinction between social and scientific is to be erased. The relativist position holds that what is 'true' and 'real' varies from place to place and from time to time. Commitment to relativism permits sociological analysis of what were previously seen as 'true' facts - ie, scientific facts, because there are no true facts independent of the social world.

1.3 SSK and the Nature of Knowledge

Within SSK a general theory about the nature of knowledge, which complements the relativistic view of the world, evolved. SSK maintains that knowledge is held within a group, and is the property of the collectivity rather than the individual. New discoveries are made and knowledge becomes acceptable through processes of social consensus. Maintaining knowledge and transferring it to others are also social processes, involving collective and consensual decisions. The idea of private knowledge, with each individual holding a store of knowledge or a bank of individualised facts is rejected. Individuals gain understanding and knowledge as they are socialised into a group and a form of life and take on board the cultural assumptions and frameworks. Socialisation, which is the process of learning to perceive the world in terms of a culture lies at the root of all understanding (Lipscombe, 1990 p. 93). We learn by being socialised into a form of life, not merely by being instructed. Collins (1985) calls this the enculturational model of learning (p. 57).

Rudwick (1985) suggests: "the new scientific knowledge produced in most episodes of scientific research practice ... should not be treated only or primarily as the creative achievement of one or a few outstanding individuals. It should be regarded rather as the outcome of processes of interaction within a group or cast list that included, in their diverse roles, not only star performers but also minor actors and walk on parts." (p. 15) (5).

This idea, that scientific knowledge is produced through processes of interaction and negotiation is highlighted by work in SSK which deals with scientific controversies. These studies stress the role of 'social construction' in scientific discovery, and show how closure of debates about the existence or otherwise of phenomenon is achieved (6).

Collins and Pinch (1993) describe several scientific episodes in order to highlight the messy character of scientific discovery. They stress that science rarely produces clear-cut conclusions, and that acceptable 'facts' are produced through social negotiation which leads to consensus about the nature of acceptable outcomes. Mackenzie's study (1989) of missile accuracy testing stresses the importance of the 'testing' phase in the social construction of 'facts' about science and technology. He states that it is during testing that "it is decided whether artefacts 'work', how well they work and what their characteristics are" (p. 411), and that this phase has a major influence on the knowledge produced.

Other studies emphasise the importance of social practices, social interaction and negotiation in what is deemed an adequate or acceptable representation or fact in the medical world:

Pasveer (forthcoming) discusses the "rendering practices" employed by X ray workers in the early twentieth century who "set out to actively fabricate the content of the [X ray] pictures", so that they were seen to represent the organ or the body. Using the lung and pulmonary tuberculosis as an example she describes how a shadow on a photographic plate was made to relate to the object transradiated, and how the precise clinical content of an X ray was determined by the *creation* of likenesses between X ray shadows and ascultated sounds and dissected lungs. She suggests that "an ever increasing self-sufficiency of the shadow pictures" emerged as a result of the practices and procedures employed by X ray workers. Brante and Hallberg (1991) use a different medical example - the 'concept of death' debate - to highlight the social interaction which closed the controversy and lead to the construction and adoption of a new definition of 'death' in Sweden in 1988. In a similar medical vein, Anderson (1992) suggests that the decision whether or not to use computerised diagnostics at a Melbourne hospital, was reached through a process of social negotiation - the stronger group in the hospital emerging triumphant and rejecting the technology (see chapter 4).

The influence of social constructivist thinking is widespread. The relevance of these kinds of analysis to this study is to highlight the importance of social processes of negotiation to the acceptance or rejection of something - whether it is a fact or a piece of technology. This will be fully discussed in my conclusion.

2. METHODOLOGY IN SSK

The debate in SSK now concerns the most appropriate way to *do* sociology of scientific knowledge. Which methodology is best? Many alternative views have surfaced, and can be overwhelming for newcomers to the field (Collins, 1983). Here I will describe the original position as set out by Bloor (1976), and some contemporary variations on his theme.

2.1 Bloor and The Strong Program

David Bloor (1976) took knowledge to mean *all* the collective beliefs that people hold and live by. Bloor's 'strong program' for SSK, which is an amalgamation of older ideas about how sociologists should approach the

explanation of scientific beliefs, stresses that an account of the nature of scientific knowledge is best achieved using the scientific method itself. Bloor (1976) set out four tenets which SSK should adhere to:
(1) SSK should be concerned with causality, and concentrate on the social and other conditions which cause all beliefs or states of knowledge.

(2) SSK must be impartial with respect to divisions between 'true' and 'false' beliefs, rational or irrational beliefs and successful or unsuccessful beliefs. By explaining both sides of these dichotomies SSK will illustrate how the perception of something as true or false is part of the phenomenon to be studied.

(3) SSK must be symmetrical and use the same types of explanation for true and false beliefs.

(4) SSK must be reflexive and its explanations must apply to sociology itself.

The range of methodological positions in contemporary SSK is based largely on the extent to which Bloor's four tenets are supported. These are summarised below (7).

2.2 The Edinburgh School

Barnes (1974) approaches the study of scientific knowledge in a similar way to Bloor, acknowledging the social aspects of scientific knowledge and rejecting the received view of science. Barnes uses an instrumental account of knowledge production, claiming that "the process whereby knowledge is evaluated involves continuing reference to shared goals and interests" (1983, p. 44). He stresses that scientific communities and institutions have a vested interest in their own perpetuation, and these interests influence the production of what is deemed to be 'scientific knowledge'. So, science is theoretical knowledge that is subject to social influences. The position occupied by Bloor and Barnes is commonly referred to in the literature as the Edinburgh school.

2.3 The Empirical Program of Relativism

The empirical relativist position is exemplified by Collins. This programme ideally consist of three stages. Collins (1981b) outlines these: The first stage is the identification of local interpretative flexibility of science, which prevents experimentation from being decisive. The second stage of the program is the description of the devices used to limit interpretative flexibility, and thus ensure closure of controversial debates. The third stage in the program will be an attempt to relate the constraining mechanisms to the wider social and political structure. Collins sets out his ideas about how to go about doing SSK within these guide-lines (Collins 1981a). He accepts parts of the strong program - but suggests that tenets 1 and 4 - causality and reflexivity - detract from the main thrust of the program. He is especially critical of reflexivity which "can lead to paralyzing difficulties" (p. 215). He does not see reflexivity as an issue for sociologists of scientific knowledge - it should be assigned to a sociologist of sociologists of scientific knowledge. Reflexivity only creates a 'regress of sociologies'. Collins suggests that "the natural world needs to be approached in a relativistic way, but this does *not* imply that the social world be approached in this way." (1981a, p. 216) His prescription is to "treat the social world as real and as something about which we can have sound data." (1981a, p. 216-217) He thus dismisses reflexivity, his line being, I am doing sociology, anyone can do sociology on me, but I do not have to do sociology on myself.

For Collins the vital parts of Bloor's framework are symmetry and impartiality. He fuses these two concepts together to form his 'Radical Program'. The radical program demands that no decision of what is a true or a false belief is made in advance, and terms such as true, rational, successful and progressive are not used in explanations. The investigators' views on such things are suspended, in true relativist fashion. They are urged to think like sceptics in order to achieve a strangeness that allows them to treat all views in the same way.

More recently, Collins has argued that relativism does not accept any epistemological stance as correct. Extending Berger (1963), his methodological relativism allows 'meta-alternation' between positions according to the purpose at hand. In 1992 Collins and Yearley were still involved in this debate, insisting that scientists should be naive realists and sociologists should be social realists. Sociologists should choose the best epistemology for the task at hand. This is the best way to sort out the relationships in the knowledge producing world, and assess science for what it is. For theorists who stress the importance of reflexivity, Collins's position is problematic:

2.4 Discourse Analysis and New Literary Forms

Woolgar and Ashmore are two prominent supporters of the reflexive line. They state (1988) that one of the first moves in the reflexive direction within sociology was Discourse Analysis, and that prior to this, reflexivity had been treated as inherent but uninteresting by Barnes and Bloor, and had been actively opposed by Collins. Discourse analysts concentrate on analysing what people say rather than what they do. Those involved (eg. Gilbert and Mulkay 1984; Mulkay 1985) seek to illuminate the processes used by sociologists to construct sociology from interview data. This disclosure of the representations used by sociologists was designed to encourage reflexive self-awareness (8).

Of late, discourse analysis has been superseded within sociology (although it is still popular in psychology) and other reflexivists have attempted to deconstruct the methods of SSK to a greater extent: 'New literary forms' advocate multi-authorship in texts so as to avoid authority and misrepresentation. This method presents all sides of any argument as though from the viewpoint of an outsider. The suggestion is that new literary forms will lead to new means of exploring old questions of knowledge and epistemology.

An example of the 'new literary form' of discourse can be found in the concluding chapter of Ashmore, Mulkay and Pinch's book (1989). This book analyses the attempt at a practical application of a social science to a specific area of social life (in this case the application of health economics to NHS administrative problems). The book concentrates on the "paradoxical and multivocal" nature of the social world and the "complex and constantly changing concatenation of voices and versions" that constitute it (p. 190). The authors recognise that social science has to operate in this social world, but that neither sociologists nor health economists can offer more that a partial account of the processes under study because of the "interpretative multiplicity of the social world" (p. 7). In keeping with this view, multivocal conclusions to the book are presented - an economists', a sociologists' and a lay person's reading of their text are set out, and through these the authors show the variations of accounts of one phenomenon that are available.

This analysis of the formal methods used by health economists attempting to solve practical problems in the NHS pinpoints three specific strategies that health economists use in an attempt to influence the inherently 'irrational' processes of the NHS. These are (i) educative strategies which replace common sense ideas with elegant economic principles (see also Mulkay, Pinch and Ashmore 1987), (ii) direct intervention in the NHS and (iii) participation by health economists in public debate. Ashmore et al highlight specific formal techniques used by health economists in the 1980s to "raise the level of operational rationality in the NHS" (p. 197) - Quality of Life measurements, Clinical budgeting and Option Appraisal. In chapter 10, I compare their criticisms of such formal methods of problem solving with may own critique of the formal methods underlying AI.

More radical reflexivists such as Woolgar press the reflexive point still further than discourse analysts:

2.5 Woolgar and Reflexivity

Woolgar stresses how vital reflexivity is for SSK (eg. Woolgar 1988). He is critical of Bloor, on the grounds that the strong programme is contaminated by essentialism, and that the concept of rules for sociological analysis is unacceptable. Rules are merely post-hoc justifications and rationalisations of action, so why should they lead to a particular kind of sociology? More crucially, Woolgar is sceptical of Bloor's description of the strong program as 'scientific', as this assumes the nature of the scientific method before it is known what that method is or whether it exists. Woolgar takes the nominalist line - science is undefinable.

The crux of the issue is in Woolgar's culmination of the discussion of representation. Representation is "the means by which we generate images of the object 'out there'" (Woolgar 1988, p. 30) and this is at the heart of essentialism. Representation sustains science and the attempts of others to study science. We can never be sure if a representation adequately reflects an object and he suggests that the problems of the object/representation connection are unavoidable in the long term, even if they can be managed in the short term.

(1988, p. 32/33 - The methodological horrors of indexicality, inconcludability and reflexivity). Woolgar stresses that "A critical appraisal of the idea of science must challenge the very idea of representation. In particular, we need to realise the extent to which our own efforts [as social scientists] are themselves beholden to the ideology of representation" (p. 36).

According to Woolgar, even progressive sociological accounts of scientific knowledge largely fail to take the bull by the horns and challenge representation. So far, all that has been done is that scientific representation has been replaced by sociological and literary representation. Woolgar's view is that SSK has not gone far enough. A reflexive exploration of sociologists' own practices and a challenge to their representations is now necessary, so that SSK can approach deeper questions about knowledge production. Woolgar and Ashmore (1988) sum this up, saying "the growing confidence with which scholars have argued that *natural* science is a social construct is now accompanied by growing interest in the consequences of applying this same argument to knowledge generated by the *social* sciences" (p. 1) (9).

The reflexivists stress the need for equivalence in treatment of work done by sociologists and those whom they investigate - scientists. Woolgar's answer to the problem of representation is to reverse the traditional idea that an object exists prior to its representation. He asks us to consider the idea that objects are *constituted* through their discovery and representation - not merely revealed by their discovery. Inverting the traditional relationship neatly disposes of the essentialist idea that objects exist out there to be discovered. Woolgar applies the inversion idea to the relationship between scientific knowledge and the natural world. In place of the belief that scientific knowledge arises from the natural world, he stresses the alternative, that the natural world arises from scientific knowledge.

2.6 The Actant Networkers

Aside from the relativist and reflexive stances, another philosophically radical branch of theory has emerged: Callon and Latour (1992), have presented the Actant Network theory. They propose a radical extension of symmetry beyond Bloor's usage. Not only should true and false beliefs be treated equally, but all dichotomies should be given the symmetrical treatment. They challenge the notion of a distinction between nature and society (as referred to by Collins and Yearley, 1992), because the divide was created by us and so ought to be disregarded. Instead they advocate treating human and non-human 'actants' in science production in the same way, and designating agency to both, in order to illustrate the co-operation of society and nature. Humans are thus removed from the central position they occupied in Bloor's version of symmetry. Latour and Woolgar used this idea in their 1979 study when they assigned agency to scientists' inscription devices. More recently, the French inspired version of radical symmetry draws no distinction between objects that have been created and those occurring naturally. Eg, Callon (1986) grants agency and power to scallops, Latour and Johnson (1988) to door closers. In order to assess the power of non-human actants, Latour suggests that we use a counterfactual method, and imagine the situation if the non-human actant was not present. We can then measure the complicity of non-human actants (10).

In contemporary SSK the debate over epistemology remains as lively as ever, yet seems inconclusive. A divide exists between supporters of reflexivity and radical symmetry on one side and relativists of the Collins/Yearley mode on the other.

3. DISCUSSION

It is not clear whether reflexivity enhances the quality and insight of sociological investigation. Challenging representation is tantamount to challenging the foundations of our taken-for-granted way of carrying on. Representation is everywhere.

Language and communication are based on common understandings of representations of things in the world. Every time Woolgar indulges in written work, is he not taking on board an understanding of the representations of the readers of his texts? Giving every viewpoint an airing so as to diminish the authority of a single author - as the new literary form of writing does - is non productive. No conclusions are reached and with every voice under the control of one writer, distortion is inevitable. Perhaps we cannot leave representation behind.

Reflexivity and critical analysis of the methods of sociology seem to be an ever decreasing circle. Ignoring the problem of sociologists' reliance on representation, as Collins does, will not solve it. Neither, it seems, does confronting the problem. Collins's belief in compartmentalising SSK separately from science, and treating the two disciplines differently, causes reflexivists much distress. Exactly what does he mean when he says that the social world is something about which we have sound data? And if reflexivity is, in effect, a regress into sociologies of sociologies that leads nowhere, why is the initial investigation of scientific knowledge of interest to sociologists?

The various positions and hotly defended epistemological stances are seemingly irreconcilable. The self-defeating nature of reflexive analysis is easy to see. But so too is the desire amongst reflexive thinkers to utilise their preferred methods in an analysis of themselves.

4. WHAT THIS MEANS FOR SSK IN RELATION TO AI

The link between SSK and AI is not immediately obvious. The notion of a 'sociology of machines' can appear counter-intuitive (Woolgar 1985, p. 558), and the justification for broadening the scope of sociology so that it includes AI and machines, needs to be clarified.

The AI phenomenon is important to sociology for three distinct reasons. Two of these were set out by Woolgar (1985):

Firstly, it offers the opportunity to extend the scope of SSK by applying it directly to the sociology of machines.

Secondly, it offers an opportunity for assessing one of the basic cornerstones of SSK - that knowledge is a social possession, held by a group rather than an individual, and which is obtained as a by-product of immersion into a form of life. AI puts forward the idea of an isolated form of knowledge and reasoning in an unsocialised machine. How can an isolated mechanical artefact access human understandings? It does not have the benefit of socialisation, or the accumulated cultural assumptions and frameworks that socialised humans do. In short, if the SSK conception of knowledge is correct, then AI cannot work. A 'sociology of machines' is a means of investigating these issues (Woolgar, 1985).

The third reason for looking at AI from a sociological stand-point has emerged since Woolgar put forward his reasons for pursuing the field. Clearly the boundaries of symmetry in SSK have been pushed back by the reflexive stream of thought. The culmination of the actant network approach is the idea that non-human actants be granted authority and agency. In essence, this symmetries-away the distinction between human and non-human. This is a boost to the proponents of AI who have long since implied that the distinction can be transcended. Looking at machines and computers will illustrate the feasibility or otherwise of erasing the distinction.

Beyond this, the field studies will illustrate the processes by which machines and computers participate in the form of life within medical establishments. Socialisation and enculturation are presented by SSK as essential to mutual understanding and integration into social groups. Does experience with medical machines in medical environments cast doubt over this fundamental claim in SSK?

A study of the practicalities of AI machines in action will contribute directly to the SSK debates, and hopefully inject new life into the discussion. Many of the philosophical and sociological questions about the nature and location of human knowledge can be illuminated by a study of Artificial Intelligence. Collins summed up this idea in 1985 when he said "the pigeons of philosophical scepticism and phenomenology are quietly coming home to roost in the nest of Artificial Intelligence." (p. 20).

FOOTNOTES TO CHAPTER 1

1. Although at this time they confined their investigations to the relations between scientists, and were, in effect, 'sociologists of scientists'. (Woolgar, 1985) The nature of the knowledge produced remained beyond the scope of their sociological investigations.

2. Merton's work spans forty years of sociology of science. His 1952 essay outlines why the sociology of science took so long to emerge from traditional sociology of knowledge. This essay appears in a 1972 book of collected pieces written by Merton from 1937 onwards. The book, edited by Storer, offers a wideranging account of the sociology of science.

3. SSK is often upheld as the empirical application of Wittgenstein's later philosophy.

4. Some theorists have feared that the intention of SSK is to change science rather than to illuminate the nature of science. This view is still held by some - eg, see Wolpert, 1992.

5. Gooding (1990) also stresses social aspects of scientific practice that are often neglected: His account concentrates on the agency of observers, how their observations are influened by interactions with one another, and he highlights the overlap of social and cognitive elements of scientific work (p. xii-xiii).

6. Eg, see the special edition of Social Studies of Science, February 1981, volume 11 no. 1. Here, work by Travis, Collins, Pickering, Harvey and Pinch clarifies the social constructionists' perspective.

7. The most recent analysis of various positions can be found in Pickering (ed) 1992.

8. According to the relativist camp, represented here by Collins and Yearley, the problem with discourse analysis was that it saw itself as invulnerable to its own critique, when in fact all knowledge making is vulnerable to the methods of SSK (Collins and Yearley, 1992, p. 304).

9. Collins and Yearley (1992) also criticise this stance. Reflexive thought highlights the omnipresence of the problem of representation. But as this problem cannot be solved, what use is there in highlighting it? Reflexivity is not a direct route to truth about the social world - just as SSK is not a direct route to nature. They support meta-alternation and compartmentalisation as a more suitable approach.

10. Collins and Yearley's response to this is that blurring the social/natural divide in this way ignores advances that have been made in the sociological study of artificial intelligence which draw attention to the difference between social things and natural things. Furthermore, resorting to empirical scientific accounts to measure the complicity of non-human actors goes against the grain of contemporary SSK (Collins and Yearley, 1992, p. 321).

SECTION 1

CHAPTER 2

THE DEVELOPMENT OF AI AND ITS APPLICATION TO MEDICAL PROBLEMS

CHAPTER 2 - THE DEVELOPMENT OF AI AND ITS APPLICATION TO MEDICAL PROBLEMS

1.INTRODUCTION

The field of Artificial Intelligence is confused by different definitions of the term AI, which arise from the different objectives and philosophical foundations of the researchers and writers involved. For example, Searle (1981) draws a distinction between weak AI and strong AI (1), whereas Clark (1989) suggests that conventional wisdom has divided the AI field into technological AI and psychological AI (also called cognitive science) (2).

In this chapter I will describe different approaches that have been adopted by AI researchers, so as to sketch out the history of AI and its progress toward the goal of making machines smarter (Winston and Prendergast, 1984). The introduction of AI techniques to the medical field is then described and the various reactions to these introductions are set out.

1.1 The Evolution of Artificial Intelligence

Dreyfus and Dreyfus (1984) suggest that the roots of AI can be traced to the time of Plato and Aristotle. More recently Alan Turing is generally regarded as the father of modern computer design - van Rijsbergen (1985, p. 282) says that Turing (1936) was the first to define on paper what we now know as a general purpose computer. The first working electronic computers appeared after the second world war. Subsequently, branches of research within AI evolved as the field matured. Dreyfus and Dreyfus

(1988) relate the history of the field since the mid 1950s using three distinct historical periods as the framework for their review:

1.1i 1955-1965 Connectionism and Symbolic AI

The first period ran from 1955 until 1965, and saw two research directions emerging: The first of these modelled the 'hardware' of the brain, simulating the interactions between neurones. This 'connectionist' approach drew on neuroscience rather than philosophy. One of the earliest proponents was Rosenblatt. His framework was based on a holistic view of knowledge and the world, and the aim was to automate the procedures by which networks of neurones in the brain learn to discriminate patterns and respond to them. His efforts concentrated on producing a 'neural network' device called Perceptron. The early successes of Perceptron led Rosenblatt on to suggest that connectionist or 'parallel distributed processing' (PDP) systems can have original ideas, generate their own abilities and learn (see Rosenblatt, 1962).

The other research approach was based on a view of a computer as a system for representing the world and manipulating mental symbols. Solving problems was the aim, as this was seen as the basis of intelligence. This approach grew from a rationalist tradition in philosophy based on the idea that manipulating symbols by rules will produce intelligent behaviour. The most famous workers in this branch of early AI were Newell and Simon who produced the General Problem Solver (GPS) program in 1956.

Connectionism and symbolic information processing both flourished initially. These approaches fell into Clark's category of cognitive science -

both aimed to produce systems to do human tasks, and they both aimed to model the psychological processes of the mind. The differences arose because each camp had different ideas about the nature of those human psychological states and processes, and how the mind works. Neural nets needed to do a great deal of computing in order to solve even simple problems, and because the available computing power was limited, it was only possible to speculate on their usefulness. Symbolic AI, on the other hand, appeared to be solving useful problems and by the early 1970s symbolic AI was at the forefront of research, overshadowing the connectionist approach (3).

However, the problems inherent in this model were soon realised: Human problem solving is not based on a few straightforward principles and rules, but is more complex and context related. Dreyfus (1972) suggests that symbolic AI was based on four mistaken assumptions. First, that biologically, the brain operates via a computer-like system of on/off switches, and that the mind processes information according to rules. Second, that all knowledge can be articulated and formulated as rules. Third, that the world can be analysed in terms of context-free discrete pieces, and finally, that background can be represented explicitly. He calls these assumptions the psychological, epistemological, ontological and metaphysical assumptions. They have survived in modern western culture because of a widespread adherence to rationalistic Platonic philosophy, that reduces all reasoning to rules and the world to atomic facts.

The question of whether humans operate according to a set of rules is a vexed one. Ideas in SSK suggest that the basis of human understanding and action is more complex than this. Predicting a rule for how a human will react is not possible in an open system because such a wide range of options are open. We can retrospectively suggest the rule that guided any action, but this is not necessarily be the rule that did guide it. It certainly is not a description of how the future performance of the same act will proceed. That people do follow rules is clear when they break a rule - but we cannot articulate those rules. Rules do not contain the rules of their own application, so it may be that the apparently rule breaking incident was the application of a new rule. The rules of human conduct cannot be specified in advance because human conduct proceeds according to mutual understandings of each other, of the world and of the context and circumstance of problems. This is all gained through shared social foundations and participation in the socialisation process. Suchman (1987) stresses these issues, and says that "The coherence of situated action is tied in essential ways not to individual predispositions on conventional rules but to local interactions contingent on actor's particular circumstances." (p. 27/28)

1.1ii 1965-1975 The Microworld and the Real World

The second development period, 1965-1975, was characterised by attempts to represent the wider context and circumstances of problems into programs. 'Microworlds' (eg, Winograd 1972) were centre stage. Microworlds are sub-sets of the real world, where everything is specified in advance. The idea was to move from the particulars of the microworld to the general level of the wider world. But this proved more difficult than had been anticipated. Essential aspects of the real world, against which human activities gain meaning, were missing from microworlds. In Haugeland's terms (1985), the questions of AI were eliminated in microworlds, because the domains were stripped of anything requiring wit or understanding (p. 190).

According to Dreyfus and Dreyfus (1988) "Cognitive simulation and microworlds were characterised by an attempt to avoid the problem of commonsense knowledge by seeing how much could be done with as little knowledge as possible." (p. 32). If the field was to progress from this juncture the problem of 'commonsense knowledge' had to be approached. The base of meanings available to people has to be made available to the program. It was felt that this approach, rather than more rules, should enable the programs to approach changing situations as people do.

1.1iii 1975 onwards - Commonsense and Background

During the third stage the AI workers have concerned themselves with the commonsense knowledge problem. 'Scripts' (Schank and Abelson, 1977) and 'Frames' (Minsky, 1974) were the first developments in this phase. Schank and Riesbeck (1981) offer a description of these kind of programs: the designers were attempting to incorporate information about background features into the programs, using a data structure with various slots which represented various features and interactions of the everyday world. A script is filled with objects, rules for how the objects relate and possible actions that can be taken by objects. The relationships represented between the objects are stereotypical.

But human life isn't quite like this. The scripts and frames were more shallow than human understandings are. Furthermore, problems of incorporating every eventuality into a script or frame were overwhelming. The researchers simply could not anticipate all possible scenes and scenarios in advance. Attention turned next to expert systems.

1.2 Expert Systems

Expert systems are comparable to microworlds and scripts and frames, insofar as they are designed to operate in sub-sets of the world. Expert systems are built on the premise that heuristic searches will be more effective if specific knowledge about the domain is made available to the program. This knowledge, when held by people, is based on accessible information about the domain and on practical experience gained within the domain. The task facing the designers of expert systems, or IKBS - 'Intelligent Knowledge Based Systems'(4), is to access the specific knowledge and incorporate it into programs. This job of "mining those jewels of information" (Feigenbaum and McCorduck, 1984) from the human experts, falls to knowledge engineers.

1.2i Definitions of Expert Systems

The definition of an expert system is not universally agreed upon. Collins gives a broad two-part definition: An expert system is a computer program that is based on the knowledge of human experts and designed to replace those experts in social interactions (Collins 1987, 1990).

Lipscombe (1990) offers further enlightenment, explaining that expert systems have three characteristics that distinguish them from other AI programs. These are:

1) Utility - they are conceived and designed to be useful to humans as tools.
2) Performance - they are designed to achieve high levels of practical performance.

3) They are transparent, which means that their rules bases can be understood by people who are not computer literate.

Madsen et al (1991) give a broader definition of expert systems as "computer systems which can evaluate information and make decisions in a manner simulating a human expert." (p. 121/122) (5). Another definition is given by Winston and Prendergast (1984), who say that "An expert system is a computer program that behaves like a human expert in some useful ways." (p. 6).

The argument for expert systems is that computerising human knowledge means that the expert system programs will display reasoning based on knowledge which is superior to reasoning based on more formal methods, in three ways:

(1) Most solutions to real problems have a social rather than a mathematical basis,
(2) Humans reach their level of performance via knowledge, so machines should mimic this technique,
(3) Codifying and distributing knowledge to a wider audience is a useful endeavour in itself (see Lipscombe 1990, p. 41).

1.2ii How Widespread are Expert Systems?

Reports of the extent of the development and deployment of expert systems differ greatly. Discrepancies in the literature are widespread. For example, in a paper in Science (1983), Duda and Shortliffe stated that only four expert systems were in regular use. In contrast to this Woolgar (1988) points to the inherent optimism of reports such as that by Resnick et al (1982) who suggested that nearly fifty expert systems had been built. Buchanan (1986), a keen supporter of expert systems, reckons that about sixty systems had been moved from development laboratories to field trials and routine use by 1986. He suggested that the possibilities for AI are endless and that the working applications - ie, expert systems applications - are continually expanding.

In the next section I explain the history of AI and expert system applications in medicine, and describe the various reactions to these applications.

2. INTRODUCING COMPUTERS INTO MEDICINE

Originally, computers in medicine were treated as general information processing devices, useful for storage and retrieval of data. Gradually, suggestions that the statistical and mathematical prowess of computers should also be applied to the practical aspects of medicine and diagnosis was raised (eg. Fox and Alvey, 1983). It seems that from the 1950s some writers (eg. Meehl, 1954) considered statistical approaches to be superior to clinical judgement - procedures such as Bayesian analysis allowed complex decisions to be approached mathematically. Some physicians became interested in computers in the belief that doctors require assistance in some areas of clinical decision making. Using computers was seen as a way of improving the structure of medicine by making it fully formal, and of facilitating the automation of many tasks (Lipscombe, 1990).

2.1 The Perceived Need For Computers

In his 1970 paper, Schwartz declared that a shortage of physicians, their geographical maldistribution and a broad societal commitment to extending health care were problems that could not be remedied unless new approaches to health care were developed. One of these new approaches would inevitably be the use of computers that were integrated into the medical system, that could take on some of the duties of the physicians and other medical staff (p. 1257).

Other writers recognised that the continued growth of medical knowledge meant that doctors could not be expected to remember the sheer quantity of relationships and concepts involved in clinical medicine. Their performance can never be 100% accurate: de Dombal (1979) suggests that because people view doctors as decision makers and healers, expectations of doctors capabilities are usually far too high. de Dombal explains that doctors use heuristics, algorithms and a pay-off-versus-error consideration when making decisions. They acquire information, analyse that information and make management decisions - how well they do these things depends on their experience, and their ability to decide whether a patient is telling the truth. The doctor allocates a patient to a problem class and discriminates amongst the diseases in that class in order to make a diagnosis. This is a complicated procedure, and not all doctors are good at all stages: All in all, an incoherent approach is evident, and a need for computerised assistance for clinicians is evident. Many researchers in both the computing and the clinical fields believed that computers which could do diagnosis would be the answer to these problems: "When the electronic digital computer first appeared it was widely supposed that one of the first and easiest tasks in medicine would be to have the computer do diagnosis" (de Dombal, 1979, p. 37).

2.2 Early Methods Adopted

In an analysis of the history of AI in medicine, and of how machines have moved into the medical environment, Shortliffe, Buchanan and Fiegenbaum (1979) outline seven major ways in which computers have been used in medical problem solving: Clinical algorithms, data bank analysis, mathematical models of physical processes, statistical pattern matching techniques, Bayesian statistical approaches, decision theory approaches and finally symbolic reasoning approaches. Characterising this progression is a shift is from pure observational data to the use of high level symbolic knowledge - computers eventually producing judgemental as well as numerical advice. The last paradigm - symbolic reasoning techniques - uses qualitative judgements and inference techniques as well as statistical methods. The attention of the program is focussed on vital bits of the problem, by heuristics which reflect practical knowledge. This is believed to be superior to decision analysis because real (human) clinical judgement is also based on knowledge, experience and rules of thumb. The move toward the adoption of symbolic reasoning shows that early medical systems, based solely on mathematical and statistical techniques and branching logic and rules, were not integrated into general medical use on a large scale (Lipscombe, 1990). The shift to more qualitative techniques mirrors the developments in AI generally, where rule based programs were replaced by microworlds and then expert systems, in an effort to incorporate more background knowledge and understandings into programs.

2.3 The Move To Symbolic Approaches

Gorry's paper (1973) explains in detail why, in medical AI, the progression from statistical approaches to symbolic methods was made. He describes a statistics based computer programme designed to act as a consultant. The program was not designed to work in the same way as a human, but used mathematical techniques and decision analysis. Gorry explains that "Before a computer can be used to significant advantage in analysing 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 formalised 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" (p. 46). The area of Renal failure was finally chosen as a suitable region in which to test this machine. The results were impressive, the machine duplicating the decisions of the experts in over 90% of the cases presented to it.

Yet Gorry was dissatisfied with the approach. He recognised that the trial was biased in favour of the machine, in such a way that it could achieve high accuracy by undertaking a 'search' of its data base. The rigid disease definitions, the number of diseases and the types of tests available all fitted into a search technique, that was amenable to a computational approach. In real situations doctors avoid large searches by using experience-based heuristics. The machine could not deal with the complexity of real cases, for example where two diseases presented simultaneously. These cases are not as sharply defined as the ones chosen for the trial. These shortcomings could be rectified by minor changes to the program, but how many minor changes would be necessary, and how could the areas in which changes needed to be made, be identified? If the machine was to be of any practical use a major change in the direction of the research was necessary.

Gorry advocated a shift to systems that incorporate 'concepts' of disease in their programs. Concepts are defined as the central, problem-specific ideas in terms of which experts organise their knowledge. He recognised that concepts may be numerous and that each one may be based on assumptions of enormous knowledge about the world. However, he thought it unnecessary for the program to have extensive knowledge about the real world, because of the precise nature of medical language in which the knowledge of the program could be expressed. In conclusion, he acknowledged that there will be cases where the program will lack knowledge relevant to a particular clinical situation, and as such it should offer suggestions to the clinician, rather than make pronouncements. Gorry's paper is important in advocating a shift to techniques involving practical knowledge of diseases.

Schwartz, Patil and Szolovits (1987) also catalogue the approaches used in AI in medicine since the 1970s. They explain the problems of pure rule-based systems and pattern matching systems, and the progression to the point where pathophysiologic reasoning was incorporated into programs. They recognise that adding this knowledge will improve the performance of systems, but will also increase the computational task. The way forward is presented as a combination of new methods of incorporating pathophysiologic knowledge with some of the older methods that have been largely discarded. Schwartz, Patil and Szolovits realised that major technical and intellectual problems still had to be solved before reliable programs would be produced. But their prediction was that by the year 2000 a range of programs would be available that would assist physicians (1987, p. 687).

In the same way, Szolovits, Patil and Schwartz (1988) pointed out the limitations of programs that rely on mainly mathematical and statistical techniques. They also advocated the use of new AI type programs that, above all, could explain how conclusions have been reached. The authors saw this as critically important for the time "when expert systems become available for day-to-day use" (p. 85).

The early confidence in symbol manipulation procedures which processed discrete pieces of information according to rules gradually waned. It was realised that the notion of representing facts about the world by symbols, and the relationships between symbols by rules in order to produce a useful system, was ill founded. The influence of early AI workers like Newell and Simon decreased as the shift to symbolic reasoning techniques occurred. The next move was to expert systems which incorporated rules of thumb and pathophysiological reasoning in programs. Since the earliest days of expert system development the medical arena has been promoted as their main area of application. The alleged achievements of expert systems have been widely publicised - eg. Wright and Bourne (1988) pointed out the successes saying that "The most successful applications ... have been in chemical and geological data analysis, computer system configuration, structural engineering and medical diagnosis." (p. 19). Similarly, Feigenbaum and McCorduck (1984) stated that "Expert systems have demonstrated that a computer is capable of the same kinds of intelligent behaviour as a physician making a diagnosis" (p. 32), and that "perhaps the largest single group of expert systems is centered in medicine." (p. 87) The medical world is a favourite site for expert system development and application. Knowledge engineers liaise with medical practitioners and encode their practical knowledge in medical expert systems. One of the most important issues associated with this process is, how much knowledge do the programs need in order to function appropriately?

2.4 How Much Knowledge Do Programs Need?

In the conclusion of his paper, Gorry (1973) forecast that medical programs would probably not need detailed knowledge of the wider world. This raises the question where does medicine end and the (real) world begin? Where should the line between the two be drawn? Can a computer program make decisions about specific medical abnormalities without a foundation of knowledge about medicine generally and the wider non-medical world? Can medicine, or sub-specialities within medicine be treated as microworlds?

2.4i The Medical Funnel

Blois (1980) discusses these issues in an illuminating paper. He utilises the concept of a funnel to represent the processes occurring between the first doctor-patient encounter and the formulation of a working diagnosis. The first stages in this process fall into the wide end of the funnel - position A. This is when the doctor first encounters the patient and the range of possible diagnoses is greatest. At this point the span of comprehension that the physician must exercise to make sense of the situation is very broad. Medical judgement, an acquaintance with the world and a grasp of commonsense issues are all required. "Clinical judgement counts for little unless it rests on a firm base of ordinary human judgement" (p. 193), states Blois. The field of possibilities open to the doctor becomes progressively smaller as the patient-doctor interaction continues and more information becomes available. The doctor narrows down the range of potential disorders, and the funnel narrows to a point, which Blois calls point B.

Blois argues that the most demanding part of the process is at point A, where the range of possibilities open to the doctor is greatest, and where a wide range of cognitive skills are required to select relevant facts and narrow down the domain. When the problem has been sufficiently sharpened by this process, so that it lies in the narrow end of the funnel near region B, different skills are needed to complete the diagnosis. Expert knowledge or a mere calculation may do the trick and finish the job.

Blois suggests that all medical tasks can be performed using judgement and some can be performed using computation. The parts which can be done by computation are at point B in the funnel. This is where computers will perform well. But the performance of computers decreases toward point A, since the necessity for wider cognitive prowess increases as the width of the funnel increases. In summary, Blois says that "a great deal of human information processing must take place before the job is turned over to the computer" (1980, p. 194). If a computer were to be utilised at the beginning of the process the programmer would have to find a way of representing a whole wider world of possibilities, which is not at present a feasible proposition. The nature of the situations encountered at point A where the whole world has to be confronted is different to the situation at point B where the human structuring of the problem has already taken place.

2.4ii Misunderstanding Diagnosis

Presumably, Gorry, in saying that the program need have little knowledge of the wider world is arguing the feasibility of computers that can perform near point B. This is not a full performance of the process of diagnosis, but is merely the performance of a small part of that overall task. As Blois pointed out (1986) "Making a diagnosis and designing or running a diagnostic program are different kinds of things" (p. 225). Blois (1986) also explains how the term 'diagnosis' has been misunderstood by the developers of early diagnostic programs. The common view is of diagnosis as a single cognitive act, comprised of making a selection between clear cut alternatives. Indeed, in some instances it can be this simple. However, physicians use the term to refer to other different processes - diagnosis may mean the process of distinguishing between classes of disease, it may be the initial starting point for symptomatic treatment, or it may be the abstract process of choosing a disease that best fits the individual disease attributes identified. Diagnosis can be highly complex, whereas diagnostic programs perform limited tasks. Blois stressed that "In the case of diagnosis and diagnostic programs I would suggest that a far too simplistic view of what physicians mean by diagnosis has been taken, and that some of the programs developed have been carried out on models of diagnosis which are overly simple" (p. 227). As a result, most early programs were written on the assumption that the program will be able to perform the whole job, and that the physician has little to contribute to the diagnostic process (6).

In a reply to Blois's article, Fogel (1980) pin-points two other areas in which he says a doctor's conduct is beyond that which a computer can reproduce: assessing non-verbal communication cues, and assessing the effect of human relations between the doctor and the patient, on the patient's presentation. Both areas demand the use of cognitive skills beyond the medical field, and hence lie nearer point A than point B in Blois' diagram. Barnett (1982) makes a similar point. He suggests that the greatest weakness of computer technology in medicine is that computers do not have the capacities of wisdom and understanding that a doctor has, which arise from an understanding and experience of everyday human existence - point A type faculties.

Blois displays a healthy awareness of the problems associated with the use of computers in clinical medicine. He continuously attempts to demarcate the tasks that computers can reliably be left to do, and apart from separating judgemental tasks from computational tasks, he distinguishes between low level descriptions consisting of mostly arithmetic terms, and high level descriptions involving clinical concepts and issues. He maintains that low level descriptions are more readily computerised than high level descriptions (1983, 1988). He presents science as divided into vertical levels - the disciplines nearest the bottom of the hierarchy are most amenable to simplistic mathematical description, whereas medicine, which is at the top of the hierarchy is less amenable to simple description because multiple levels of association and cross-hierarchical explanations at various scientific levels are involved. He does not envisage complex clinical concepts as suitable for a computational approach because they cross so many hierarchical boundaries. As an example he cited the DENDRAL system (eg. see Buchanan, 1986) which he said, worked well in trials because it functions in a low level domain - molecular chemistry - and works directly on low level 'hard' science data input. On the other hand INTERNIST (eg. see Miller et al, 1986) relies on the physician acting as interpreter between high level input data of a complex nature - smells and observations for example, and the input to the machine. Blois is not completely pessimistic about the value of computers in medicine and recognises the potential for computers to assist physicians in some areas of medicine (7).

The evolution of AI in medical settings and the appearance of expert systems has produced a wide range of reactions in the literature. These are summarised in section 3, and the necessity for an investigation to clarify the current state of play is then justified.

3. REACTIONS TO THESE DEVELOPMENTS

In 1970 Schwartz's comments epitomised the expectancy of the era when he stated that "Computing science will probably exert its major effects [in medicine] by augmenting and, in some cases, largely replacing the intellectual functions of the physician" (p. 1257).

Schwartz cited the use of computers for ECG analysis and patient history taking as successful examples of the type of things computers could achieve. According to Schwartz's analysis, computers would render doctors redundant in two ways - firstly by actively taking over traditional physician tasks. He went so far as to suggest that intricate jobs such as anaesthesia administration could in principle be handed over to a computer. Secondly, computers would enable less skilled paramedical staff to perform new tasks and thus encroach into the domain of the doctor. Paramedics would be guided by the computer, the in-built instructions ensuring that these less qualified staff did not exceed their capabilities. Physicians would thus be free to concentrate on the 'uniquely human' aspects of their jobs.

Maxmen (1976) takes the same stance, saying "I advocate and predict that in the 21st century doctors will be rendered obsolete by the collaboration between the computer and a new breed of health care professionals - the medic ... I maintain that the emergence of a post physician era will be feasible, desirable and inevitable." (p. viii). He goes on to explain the role of the medic: "Under this system computers would render most of the technical diagnostic and treatment decisions currently being made by physicians, while medics, a hitherto unknown type of health care professional, would provide the supportive and some of the technical tasks currently being performed by doctors." (p. 7). The upshot of these predictions would be that "Because a medic-computer symbiosis would usurp all of the tasks presently assigned to physicians, doctors would be rendered obsolete." (p. 7).

In the 1970s, Schwartz and Maxmen's ideas represented those at one end of the spectrum - the optimist's end. They saw computers taking over all manner of tasks as they moved into medicine in earnest.

Schwartz concentrated on the probable effects of introducing diagnostic computers: Problems would arise if computers were used to appraise doctors, or if computers became too advanced for junior staff to use. Centralisation of knowledge in programs could damage the market place of medical ideas, programmers may adopt an autocratic role and the emphasis in medical education would alter. Nowhere in his paper does Schwartz discuss the *feasibility* of computers doing diagnosis and other human tasks. It is taken for granted that such an advance is possible, and the revolution in health care technology is taken as given. Using computers is presented as an inevitable step. The concept of some processes in medicine being unsuitable for a computational approach is not discussed. Only the problems of incorrect programming and system breakdown due to power failure are mentioned. Schwartz's opinion at that time was that "The power of the information sciences is such that it will, without doubt alter the face of medicine and we can ill afford to ignore this impending reality" (p. 1263/4). A revolution was pending, the full effects of which would materialise by the end of the century.

In 1987, more than a decade after his first prediction, Maxmen reiterated his claim that the post physician era was inevitable (Maxmen, 1987, p. 109).

3.1 Optimism gives way to Pessimism

By 1987, though, Schwartz, in collaboration with two colleagues, seemed to be wavering: Szolovits, Patil and Schwartz (1988) were aware that systems which were designed to simulate expert reasoning had not led to clinically useful programs. These authors still expressed confidence, suggesting that programs would eventually develop into more than diagnostic implements, and be used to plan therapy. But despondency was in the air - Schwartz, Patil and Szolovits said in 1987 "After hearing for several decades that computers will soon be able to assist with difficult diagnoses, the practising physician may well wonder why the revolution has not occurred ... few, if any, programs currently have active roles as consultants to physicians" (p. 685). They recognised that "it has become increasingly apparent that major intellectual and technical problems must be solved before we can produce truly reliable consulting programs" (p. 687). By the late 1980s, the early optimism seemed had given way to a barely concealed pessimism.

Van Der Lei, (1991) adds further weight to this pessimism: "There is no evidence that the capabilities of computers will ever approach those of human beings in dealing with unexpected circumstances, in understanding patients in their social context, in integrating the often complex and confounding presentation of a disease into a coherent pattern, or in dealing with ethical issues." (p. 1508).

The initial expectancy did evaporate and the confident prediction that a "revolution" was imminent seemed to dematerialise. Potthoff et al (1988) went as far as to say that "The speculation that computers can replace the reasoning of skilled physicians in the foreseeable future is obviously highly unrealistic." (p. 125) It became clear that "The optimistic expectation of 20 years ago that computer technology would also come to play an important part in clinical decisions has not been realized, and there are few if any situations in which computers are being routinely used to assist in either medical diagnosis or the choice of therapy" (Barnett, 1982, p 493).

3.2 Renewed Interest

However, since the late 1980s, interest in research and development into the use of computers in medicine has not declined in the way that this pessimism would suggest. In 1987, the Index Medicus listed 41 entries under the heading 'Expert Systems'. This increased to 93 entries in 1989, 104 entries in 1990, 101 in 1991 and 104 entries under this heading in

1992. Interest in the field is still running high. In 1992 the Milroy Lecture, delivered to the Royal College of Physicians of London, was based on the usefulness of computerised decision analysis in medicine. The authors, Lilford and Thornton, suggest that decision analysis is "the most threatening [development of formal logic in medicine] as it seeks to replace, or at least augment, clinical judgement by means of a formal model." (p. 401). They conclude that there is space for decision analysis in medicine, in devising treatment policy, in research in medical ethics and in deciding how to distribute scarce resources. Thornton, Lilford and Johnson (1992) also promote the use of decision analysis in medicine. They implicitly suggest that some parts of medicine are amenable to representation in a computer program.

It is clear that manufacturers, computer scientists and some medical professionals are still interested in developing computerised machines for use in various areas of medicine. The failures of the past two decades, together with the apparent dearth of systems in clinical settings has not dampened this enthusiasm. The Index Medicus list of Expert Systems for the first five months of 1993 includes papers on expert systems that are designed to diagnose focal bone lesions, epileptic discharges and cutaneous melanoma. Also listed are papers dealing with the use of expert systems for the prediction of protein localisations sites in eukaryotic cells, for use in optometry and histopathology, for teaching nursing diagnosis and many papers on expert system evaluation. With interest running high, and research and development continuing, this field remains an exciting area for investigation. My interest is in machines which perform tasks that previously fell in the human domain, and which do the job in ways that are useful in that particular medical context (8).

3.3 What is 'Medicine'?

There are obviously problems associated with the notion of machines to do human jobs in medicine. These can be reduced to different beliefs about the specific nature of medical knowledge: Is medicine a series of inter-related 'facts' that are context free? Or is it affected by external (to medicine) influences, and by circumstances and patient-specific factors? Is medicine a collection of discrete indisputable facts or is it a social practice?

Opinion is divided. Engelhardt (1974), takes the latter view, concluding that "there are no simple, pure facts in medicine, but that all facts appear in and are influenced by contexts of expectation and value." (p 225) SSK would argue that medicine is a social practice, just as science is. Maxmen clearly see things differently. His belief in the imminent replacement of the physician grows from a view of medicine as completely formalisable and programmable.

3.4 Clarifying the Confusion

The field is characterised by confusion and conflicting claims: The SSK view of knowledge is challenged by computers that appear to do medical tasks. The computers have no grounding in medical socialisation. The designers have attempted to 'microworld-off' bits of medicine as required. Does this strategy work? The problems of adapting socially gained medical knowledge into computerisable format cannot have escaped the attention of the designers and the users of these devices. How do they deal with these problems? This practical study of the current situation is designed to clarify the evident confusion.

FOOTNOTES TO CHAPTER 2

1. Both are aimed at computer simulation of human cognitive capacities. For weak AI the value of the computer is as a tool for the study of the mind. In strong AI the computer is not only a tool for studying the mind, but it is also considered to have cognitive states - if a computer is given the correct program that describes mental processes, then thought will be caused. Searle concentrated on the problems of strong AI, accepting the aims of weak AI. Although this acceptance of the notion of using an artefact to illuminate the features of cognition is puzzling: Would the supporters of this view argue that building a paper flying device illuminates the mechanisms of flying? Does a flying paper plane tell us anything about the intricacies of a humming bird's abilities to fly, hover and land? Or does a model of a bridge tell us how that bridge is likely to behave in practice? The recent experience of the bay bridge in San Fransisco, as well as the inaccuracies of many weather forecasts suggests that this kind of 'scale model' approach is flawed.

2. For technological AI, the aim is to build machines that simulate the input and output profile of a person doing a task. For those working in this field within AI there is no intention to model the mind, only to mirror human performance. For psychological AI, or cognitive science, the aim is again to get a machine to do a human task *and* to do it in a way that models human performance. The aim is to build a model of psychological states or processes - a model of the mind.

3. How the 'debate' about which research program should continue was 'closed' is an interesting story. Lipscombe (1990) discusses this episode and suggests that connectionism was sabotaged by Minsky and Papert - two researchers from the opposite camp - in their book 'Perceptrons' (1969). Olazaran (1991) also examines this closure episode. He suggests that events were more complex, and that Rosenblatt's own book (1962) pointed out the problems of the connectionist approach, seven years before Minsky and Papert. So he was largely responsible for the lack of enthusiasm for his own branch of research.

4. Both names, expert systems and intelligent knowledge based systems, are misleading. 'Expert system' implies expert level performance. Whether this is achieved remains to be seen. The term expert system also exacerbates the computer's 'placebo power' (Wyatt, 1991), which can encourage inexperienced doctors to invest too much confidence in the advice of the 'expert' computer. 'IKBS' on the other hand, suggests a level of performance equitable with 'intelligent' human performance. Intelligence is an ill-defined concept, to be treated with caution.

5. It is not entirely clear whether Madsen means that the outcome simulates a human's outcome, or whether the computer method of producing the outcome simulates the methods used by a human.

6. In a similar vein, Paul Atkinson pointed out how popular public perception of diagnosis is fundamentally flawed. It is often assumed to be an individual act, a decision made by one expert qualified to decide. Research shows that diagnosis is very much an interactive process, involving consultations between experts within medicine. The outcome is often the product of informed discussion between several participants. (Personal conversation, July 1992)

7. Blois was involved in the development of the RECONSIDER project (See Blois et al, 1981). This was a computer program designed for use as a diagnostic aid rather than a diagnostic program. Its purpose was to list all diseases that a physician should think of given a particular list of clinical findings. It was not designed for making specific, full-scale, accurate diagnoses.

8. In view of this Winston and Prendergast's definition of an expert system (1984) suits my purpose.

SECTION 1

CHAPTER 3

COMPUTERISATION AND DESKILLING?

CHAPTER 3 - COMPUTERISATION AND DESKILLLING?

1. INTRODUCTION

What are the likely effects of the introduction of computers into the work place? Will they prompt the development of a workforce with new skills - programmers and fault finders? Or will a dissatisfied, displaced and deskilled pool of workers emerge? The deskilling debate is vital: If deskilling is inevitable, is the use of machines advisable? If 'tacit' skill (Polanyi, 1962) is a barrier to automation how do we explain computer achievements to date? In this chapter Braverman's theory is presented as the seminal work in the 'deskilling' camp. In contrast to this, recent work which highlights 'tacit' skill is set out. Finally, various attempts to explain the extent to which computers can replace humans are summarised.

1.1 The 'Deskilling' Theory

Braverman (1974) argues that the driving force behind technological change is the capitalist search for increased profit, production and control. Capitalism tightens control over workers by increasing mechanisation and the automation of production, thus removing their skill and initiative. Crucial to the deskilling theory is the increased 'rationalisation' of the labour process. Taylor first presented this idea (1947) in the Scientific Management Thesis. There are two aims to this thesis - first, deskill job content by fragmenting tasks into simple operations with guide-lines for the quickest mode of operation, and second, concentrate conception and planning of tasks in the hands of management. Separating conception from execution in this way gives management a monopoly of the 'brain work' and denies workers control.

The rationalisation and control that Braverman described reflect Taylor's vision.

Most of Braverman's ideas have been criticised: his description of the working class (1); his simplistic account of the implementation of scientific management (2); his suggestion that deskilling causes a 'proletarianisation' of the workforce (3) and his definitions of skilled and unskilled work (4).

2. CONTEMPORARY CRITICISMS OF BRAVERMAN

More recently, empirical studies, mostly within industrial sociology, have offered new ways of looking at the concept of 'skill'. These contemporary contributions to the deskilling debate concentrate on types and distribution of skill at the workplace and the notion of a tacit dimension to skill.

2.1 The Distribution of Skill

Jones (1990) challenges Braverman's view that computers will replace skilled human workers and cause an overall deskilling of the workforce. He points to the influence of trade union tactics which can interrupt the introduction of automation, and he draws attention to the uneven distribution of skills amongst workers: A programmer has different levels of skill and knowledge to a machinist, yet both are essential for the introduction and smooth running of computers. In Jones's case study "the programmer had to ensure that every single instruction to cover the most minute occurrence had been written into the program" (Jones, 1990, p. 6). To do this effectively, programmers consult with machinists. This cooperation is vital during programming, and when practical malfunctions have to be rectified. The different kinds of skilled workers remain essential even in highly automated factories.

2.2 Separating Conception and Execution

Separating conception from execution to isolate all the 'brain work' is a fundamental part of Braverman's theory. Jones tackles this, drawing attention to the 'brain work' required by all workers as they plan for and perform their jobs. He suggests that no jobs require zero conceptual abilities - manual jobs demand a multitude of small conceptual decisions, whilst conceptual or intellectual jobs demand a knowledge of manual abilities. Jones's analysis implies that treating execution tasks as computerisable, because they require no conceptual ability, is misguided. This is not the place to draw a line separating the tasks that computers can do from those that they cannot do.

2.3 Tacit Knowledge

Jones (1983) explains the tacit knowledge concept: Acquiring knowledge inevitably requires learning by doing. Learning through speech or instruction is insufficient as it ignores aspects which are not articulated by practitioners. This is the tacit knowledge aspect. Because tacit knowledge cannot be articulated it cannot be prospectively formulated into rules. Since computers work using articulated rules of human conduct, it follows that computers will not have access to tacit knowledge and will not display tacit skills. Here Jones is drawing on the work of Polanyi (1962) who introduced the term 'tacit knowledge'. Collins (1985) says that tacit knowledge is what we know by virtue of our participation in a form of life

(p. 77). Again, it is clear that computers are excluded from gaining tacit knowledge because they do not participate in our form of life. Automated factories are based on "A view of knowledge in which all human thought and action can be logically described in a formalized language, and in which all conceivable activities are predictable" (Gullers, 1988, p. 37). This ignores the tacit dimension and inarticulable skills of people.

Jones's work shows that workers and the tacit skills they hold are indispensable in a well run workplace. The 'unacknowledged tinkerings' and the minor modifications at the point of production are essential if the cognitive disjunctures between programmers and manual workers are to be smoothed. Braverman did not recognise that the cooperation of workers, whose jobs are supposedly automated, is essential at the workplace.

2.4 Working Knowledge, Practical knowledge, Knowledge of familiarity and Propositional knowledge

Other writers have referred to the 'tacit' dimension using different terms: Kusterer (1978) studied two groups of workers - machinists in a paper cone factory and bank clerks. He suggests that there is no such thing as an 'unskilled' job. He identifies 'working knowledge', which is divided into (1) basic working knowledge - exercised constantly with no conscious effort, and (2) supplementary working knowledge - used periodically and consciously applied to unexpected problems. Kusterer says that "what managers, social scientists, and even many workers themselves do not realise is the extent of supplementary knowledge that is also necessary" (p. 45/6). Supplementary working knowledge is the know-how, judgement, and responsibility that is a necessary 'hidden extra', required occasionally. To avoid mistakes in these jobs, "it is not enough to simply know the rules and procedures" (p. 85) - supplementary working knowledge is also essential. Kusterer concludes that successful production "depends firmly on the know-how of workers and on their willingness to use that know-how to guide their collective work" (p. 188). Only people amass supplementary working knowledge - it is not accessible to computers. This fits in with Jones's assertion (1983) that "the primary element of craft skill which has always been largely tacit knowledge can never be completely absorbed into a formal system of representation, instruction and control" (p. 9).

Kusterer's division hinges on the idea that basic working knowledge requires no conscious effort, whereas supplementary knowledge requires conscious thought and effort. The implication is that using basic working knowledge does not require 'know how' or skill, and is easier to mechanise (5).

Noble (1984) pursues a similar point. He discusses the difficulties of transforming activities into machine-readable terms: He says that anticipating every possible variation in tools, machines, machine performance or conditions cannot be done prospectively. "Of course" he says, "such contingencies are routinely dealt with by machinists and machine operators relying upon their skills and accumulated experience with just such challenges" (p. 344). This is a reference to tacit skill, which is evident in all jobs (6).

Gullers (1988) explores the same issues. He describes skill as a combination of practical knowledge, knowledge of familiarity and propositional knowledge. Propositional knowledge is the type available in books. On its own such knowledge is impotent. For understanding to arise, it must be combined with practical knowledge and knowledge of familiarity. Knowledge of familiarity is like tacit knowledge - it cannot be articulated and absorbed into a system accessible to computers. Goranzon (1988) follows the same framework in his discussion.

All of these analyses of human activity suggest that a component exists which cannot be described verbally. This component arises from experience, know-how and practice. The essential point is that this aspect of skill is involved in the execution of all jobs. Braverman's theory is that the practical part of all human tasks can be stripped of all skill as the task is divided into many small constituent parts, and conception is separated from execution. Separating conception and execution so that 'skilled' and 'unskilled' categories emerge is problematic if tacit skill is acknowledged.

2.5 Removing the Tacit Dimension

However, Rolfe (1990) makes an interesting point. She states that "jobs involving judgement and decision making are less compatible with computerisation, and jobs which involve processing large batches of similar information are more adaptable to computerised techniques and deskilling" (p. 109). It seems that for tasks involving judgement and decision making, the tacit component is vital, whereas for tasks such as 'processing large batches of similar information', the tacit dimension is less important. To a certain extent it can be 'formalised out' without loss to the performance of the task. However, because tacit knowledge is so important for most jobs, the range of tasks that computers can accomplish is small, and so "important gaps remain between the potential of [flexible manufacturing systems] and their operational reality" (Jones, 1989, p. 49).

Rolfe (1986) presents the same idea: In this study she concluded that insurance underwriting clerks were deskilled by technology and their numbers were reduced by two thirds, whereas senior underwriters were not replaced by machines. Rolfe maintains that the lower level clerks were deskilled by computerisation. But it is important to note that standardisation of the tasks and a new division of labour between lower level clerks and senior underwriters, occurred prior to the computerisation. Lower level clerks were assigned only routine policies, and a manager at the firm stated that 'zombies' could do this work, as it is 'merely a process'. Senior underwriters, on the other hand, dealt with more complex unusual cases. The different jobs were organised into these categories prior to automation. Clearly "The two processes of standardisation and computerisation are not separate; in order to achieve maximum efficiency, standardisation of the process is required" (Rolfe, 1986, p. 43). So the low level clerks were not deskilled *by* computerisation - their jobs were reorganised so that the tacit knowledge was no longer essential, and could largely be 'organised out'. This rendered the job suitable for automation. The reorganisation came prior to the automation.

Rolfe also studied Ordnance Survey draughters. Here she did not detect deskilling, or replacement of workers by computers. The nature of the job meant that tacit skills could not be 'organised-out' by a restructuring of the task. The draughting procedures were not amenable to total standardisation, and thus defied computerisation.

It emerges from this that the potential for computerisation of any task is a function of our ability to restructure and regiment it to such an extent that the tacit requirements can be removed without substantial loss to the performance of the task.

3. BACK TO BRAVERMAN

Braverman says deskilling occurs as technology advances and computerisation increases. But this discussion suggests that because computers and automatic machines have no access to tacit skill they are ill-equiped to take over many human jobs. Only in cases where work is already organised in a routine, predictable way so that the tacit skill has been removed, can a machine can be expected to perform the task adequately. An important point is that such tasks are uncommon, and this is why such things as workerless factories are rare. When a task is structured in a suitable way, a machine can be programmed to mimic the human performing the task. This idea has been explored by Collins (1990) in his attempt to set out what computers can and cannot do, whilst other theorists have looked for different ways of explaining the issue:

4. DIFFERENT WAYS OF EXPLAINING WHAT COMPUTERS CAN AND CANNOT DO

4.1 Dreyfus and The Knowledge Barrier

Dreyfus (1972) describes four types of human activity. The first two types can be described exhaustively, and are independent of situational influences. The third category is an extension of this, and incorporates self-contained problems independent of external context. Type 3 problems are of a greater magnitude, and although fully describable in theory, they are not so in practice - an often cited example is the game of chess. Dreyfus expected to see good progress in the computer simulation of areas one and two, and some advances in reproducing category three activities, because they are all 'formal' activities. Area four is of a different order. It covers all activities which are dependent on the situation and context in which they occur for their meaning. Area four covers 'informal' activity and is not suitable for computation. The barrier between the formal and the informal is the stumbling block for computers. Collins (1990) referred to this divide as the 'Knowledge Barrier'.

4.2 The Development of Human Expertise

Dreyfus and Dreyfus (1986) offer a variation on this theme. Computers should not be applied to tasks demanding expertise. They present expertise as the product of a 5-stage process (see also Dreyfus 1987, Dreyfus and Dreyfus 1984) - novice, advanced beginner, competent performer, proficient performer and expert. The model can be summed up as follows: "The novice and advanced beginner exercise no judgement, the competent performer judges by means of conscious deliberation, and those who are proficient or expert make judgements based upon their prior concrete experiences in a manner that defies explanation" (p. 36). Dreyfus and Dreyfus suggest that machines cannot function like experts because they lack involvement, holistic understanding, expertise and intuition. If this theory is to fit in with Dreyfus's '4 kinds of activity' model, expert performance must be a level 4, informal activity, dependent on situation and context. Computers should be limited to tasks requiring 'calculation' and excluded from tasks involving 'judgement' (7).

4.3 Limited Domains

Winograd and Flores (1986) suggest that building an account of human cognition in a computer program is only possible in areas where the background of the domain in question can be fully articulated (p. 75). They see some areas as well circumscribed, and some domains 'limited' enough for programs to be effective (8). Winograd and Flores see computerisation as possible within these limited domains which are thus fully specifiable. Computers would flounder if applied to wider world tasks where context and background factors are essential parts of the problem.

4.4 Collins and Types of Human Action

Collins (1990) rejects Dreyfus's knowledge barrier, insisting that dividing the world into formal and informal tasks does not explain achievements that computers have already made in the 'formal' sphere. The knowledge barrier makes formal tasks look too easy, and if a knowledge barrier did exist it would be continually moving. Collins maintains that all activity is social, shaped by social interaction and language; SSK has shown how social influences affect all aspects of the world, including science and mathematics (which would lie on the formal side of the barrier). How a computer can be successful in *any* area of human activity needs to be explained. Collins's position is that there is only one type of knowledge in the world, and any apparent division is made by us as we deal with the knowledge. The paradox of isolated, unsocialised computers that somehow perform well amongst people operating within a form of life is what Collins tries to explain.

Collins begins with the distinction between action and behaviour. Acts are guided by intentions, behaviour is unintentional (Searle, 1984). Collins extends 'behaviour' to include the physical counterpart of an action. His example is a wink and a blink. A wink is an action guided by the intention to, say, indicate friendship. A blink by contrast is an involuntary movement of the eyelid - not an action at all, merely a bit of behaviour. His extension of the term behaviour means that the physical part of the action of winking (the blink-like movement of the eyelid) is also a behaviour. The behavioral counterpart of a wink is the same as a blink. Collins explains that

"(1) the same piece of behaviour may represent many different acts.
(2) the same act may be executed or (to use computer jargon) "instantiated" by many different behaviours." (1990, p. 32)

People sometimes choose to carry out an action in the same way every time. They forego the option to instantiate it with any number of different behaviours. Collins terms this category of action 'behaviour-specific action' as opposed to the usual kind of variable action which he calls 'regular action' He explains that "Behaviour-specific acts are acts that humans always try to instantiate with the same behaviour." (p. 33). An example is the conduct of assembly line workers abiding by Taylorist guide-lines on the best way to perform a specific operation. Observers cannot tell the difference between real action and behaviour because the visible behaviour is the same each time.

Behaviour-specific acts can be fully described according to rules which capture the behavioural coordinates of that act. These rules will then apply to all future performances of the particular behaviour-specific action under consideration. This is where behaviour-specific acts differ from regular acts. When humans engage in regular action a set of rules guiding the action can be formulated after the event, but these rules do not work as predictors of how the same act will be carried out in the future - the range of instantiation options open is too broad. Behaviour-specific acts are special because people have chosen to ignore the wide range of instantiation options. However, a behaviour-specific act is still an act, because at any time the person involved could choose to execute it with a different behaviour. The crucial point is that programming a computer with the coordinates that reproduce this repetitive behaviour-specific response is sufficient for that computer to reproduce the behaviour-specific action.

Collins points out that behaviour-specific action does not come easily to people, but takes practice. But even when people fail in their attempts, the behaviour-specific acts remain behaviour-specific because the essence of a behaviour-specific act is the *intention* to perform in a behaviour-specific way.

Collins applies his theory of behaviour-specific action to some of the mental acts we carry out. In just the same way as for physical acts, it is the mental aspect of behaviour-specific acts that machines are able to carry out.

Jones (1990) distinguishes between behaviour-specific acts and encultured actions which are comprehensible only within a form of life. Just like tacit skill, encultured action cannot be captured in a set of rules: "the quality which determines whether a skill can be automated is whether it is already machine-like [ie, behaviour specific]; that is having the repetitive character of an algorithm. Encultured actions on the other hand do not have this character because the decisions and rules that make them up are only knowable by induction form the culture of a society or social group". (1990, p. 15)

To summarise: Machines can reproduce behaviour-specific actions. They cannot reproduce regular action that does not proceed according to a set of instructions. Collins is careful to emphasise the point that "The possibility of mimicking an act mechanically depends not on whether the rules of performance are interpreted self-consciously, but on whether the act has the potential to be performed in a behaviour-specific way." (p. 217). Behaviour-specific action and regular human action can both be executed self-consciously or unself-consciously. The self-conscious/unself-conscious distinction is subservient to the regular action/behaviour-specific action distinction.

5. REASSESSING 'DESKILLING'

In essence, Braverman's deskilling account is based on the notion that automation rests on a separation of conception and execution, a division of tasks into skilled and unskilled, and the subsequent deskilling of the workforce. However, the identification of tacit elements of knowledge and skill suggest that execution is not devoid of skill. Separating conception from execution is problematic because of the tacit dimension. It seems that the potential for automation of a task depends on our ability to reorganise it so that the tacit component is removed and it can be performed in a regimented or behaviour-specific manner. Rather than deskilling occurring because of automation, automation can only proceed if we deskill our performance first (9).

Some concern has been voiced about the treatment of human intelligence as something we can pour into a machine. Weizenbaum (1976) discourages this view of intelligence as a static, unalterable, culturally independent attribute, which can even be measured objectively by such things as IQ tests. He suggests we view intelligence as a multi-faceted phenomenon that manifests itself in different ways relative to specific social and cultural conditions. Intelligence is not quantifiable or measurable on a linear scale, and debates about intelligence are sterile. In view of this he says there are good practical and philosophical reasons for moving on to empirical investigations into what computers are achieving.

One of the objectives of this research is to produce this kind of empirical work. A practical investigation of what is being achieved by computerised automatic machines is more important then whether they are 'intelligent'. If computers are acceptable to the people using them, it does not matter whether or not they are intelligent. Weizenbaum's other concern, that the debate should be conducted in terms of what computers 'ought' to be doing is less relevant - computers are already installed in various medical environments, and we cannot now indulge in debates about whether they ought to be there.

FOOTNOTES TO CHAPTER 3

1. Heather Rolfe (1990) suggests that his stance is too objective because his description of working class consciousness reflects only the conditions imposed by capital. The working class is presented as passive - merely a class in itself, rather than a class for itself. As a result Braverman ignores the possibility of individual or collective worker resistance to deskilling, either at the point of production or at the societal level. The capitalist class, by contrast, is presented as homogenous and unified in its objectives. Rolfe suggests that neither of these class descriptions is broad enough.

2. Braverman implicitly accepts that Taylorism is easily installed and widely used as the most effective means of instigating managerial control. But is the logic of Taylorism really the same as the logic of capitalism? Taylorism is aimed at increasing control. Capitalism is aimed at profit rather than directly at control, and it does not necessarily follow that management will sacrifice profit in favour of gaining control.

3. Proletarianisation implies that previously the workers were something other than proletarian. Braverman is using the craftworker as his benchmark for a skilled worker. But is this notion of an autonomus artisan, overly-romantic? Cooley (1988) suggests that some writers "believe that before the industrial revolution the populace spent its time dancing around maypoles in unspoilt meadows and writing sonnets in its spare time. It was never like that" (p. 127). Similarly with Braverman's image of the craft worker. It was never like that, at least not for most workers, who were already proletarian, manual, domestic or farm workers.

4. Defining 'skilled' work and 'deskilling' is difficult. Rolfe (1990) shows that indirect measures of skill based on length of training and education are misleading. So too are direct measures of skill which use coordination of physical and mental operations in each task. Linking skill to control as Braverman did is also problematic. He implied that reducing workers control over their jobs and separating conception from execution meant that workers were deskilled. But some workers with little control can be highly skilled.

5. This is where Kusterer's analysis differs from Jones's. Jones's suggestion is that all tasks require tacit skill, and that skilled performance is often executed unselfconsciously. This would seem to be the case with highly skilled sports professionals participating in fast games. Squash players, for example, react less on the basis of conscious thought and more on the basis of unconscious reactions, perfected by hours of practice. Distinguishing between tasks that are unconsciously or consciously performed is drawing the line in the wrong place as far as the potential for automation is concerned.

6. A colourful passage in Noble's book (1984), which is a long extract from an interview with a machinist shows how subtle and invisible the 'tacit' skills of the craftsman really are (p. 344-346).

7. Dreyfus and Dreyfus acknowledge that the designers of the RECONSIDER system (see chapter 2, and Blois et al, 1981) were on the 'right track'. The role of this machine was assistant to the doctor, and it was not designed to perform 'expert' level tasks.

8. Recognising specific infections and interpreting electrocardiograms are two areas in the medical field in which Winograd and Flores predicted successful computer application, because these areas are 'limited domains'.

9. Not all commentators are interested in this particular issue of computer capability. Eg, Weizenbaum (1976) takes a moral stance on the question of computer capability, suggesting that explorations of what a computer will be able to achieve, and what the limitations may be, are misguided. More concern should be directed toward the issue of what computers should be allowed to do. The question he pursued was, "What human objectives and purposes may not be appropriately delegated to computers" (p. 207). He was not concerned with the belief in AI circles that there are no tasks that cannot in principle be handed over to a computer. Neither was he interested in theoretical debates to the contrary. Rather, he makes the point that "Ultimately, a line dividing human and machine intelligence must be drawn" (p. 8), and where this line is drawn depends on which tasks *ought* only be attempted by humans.

SECTION 1

CHAPTER 4

WHAT IS MISSING FROM THE CURRENT LITERATURE?

CHAPTER 4 - WHAT IS MISSING FROM THE CURRENT LITERATURE?

1. INTRODUCTION

In this chapter the major points and problems raised in chapters 1, 2 and 3 are summarised, the direction of the empirical research is set out and the specific questions that the research will address are presented.

1.1 The Medical Form Of Life

As doctors, nurses and paramedics undergo training they absorb the norms of the medical society and become part of the social collectivity and form of life of the hospital. SSK applies as much to medicine as it does to science - medicine is shot through with social influences - diagnosing, treating and rehabilitating patients are all social processes. The influences on the way that medicine is performed are wide ranging - physician skill, availability of facilities, contacts in the field can all affect the treatment process. In the case of each patient, the equipment, the location, the disease, its presentation, degree of severity and history, as well as the patient's circumstances and social situation all play a part. Doing medicine is a process that overlaps with being in the (wider) world. Supporters of AI suggest that these kinds of wider world details can be made explicit and formalised. For example, Slezak (1992), in his response to Collins's 1990 book, supports the notion of all-encompassing formalisation (1). Designers of computers for the performance of medical staff jobs treat all the symptoms and signs of diseases as formalisable and all the physician strategies as specifiable. Clearly these positions are at odds with those taken by Sociologists of Scientific Knowledge. Looking at medical applications focuses the wider questions facing SSK to a specific form of life.

1.2 SSK, AI, Deskilling and Medicine

Working computers pose a problem for sociologists of scientific knowledge and for industrial sociologists. For SSK the problem is, if knowledge is gained through socialisation in the human collectivity, how can AI claim to put that human knowledge into an unsocialised machine? The problem for industrial sociologists is, if tacit skill is inarticulable, unspecifiable and necessary for the execution of most tasks, how can AI claim to be able to mechanise so much of what we do?

1.3 Explanations

1.3i Behaviour-Specific Action, Digitization and Repair

The theory of behaviour-specific action (chapter 3) is Collins's answer to these questions. But Collins goes further than this in order to explain how machines fit into society. There are several strands to this argument: Collins suggests that very often machines that don't really work look like machines that do work because the gaps between a machine's performance and a person's performance are filled in by people. We contribute to the machine's level of acceptance by showing benevolence and smoothing over the machine's mistakes. This human contribution occurs at both ends of any process that is automated. At the beginning, we contribute by providing the machine with an input it can deal with. Collins terms this 'digitization'. At the end, we take the machine's output and we make it fit into the real world. This he terms 'repair'.

The concepts of digitization and repair are central to Collins's thesis: He explains that digitization is a means of dividing things up into pre-specified, exhaustive categories. No in-between categories are allowed. Within each category a degree of variation is tolerated by us, so, for example, we all accept every kind of letter Z as fitting into the Z category. The invariance of the Z category is preserved as we learn to tolerate the variations within the category. When Collins suggests that we digitize the input to a machine he means that we fit the ambiguities of the world into specified categories that correspond with the program of the machine. Digitization turns the concerted action visible in the world into pre-specified exhaustive categories of concerted behaviour. Computers can work in digitized worlds because the digitization process removes context related factors. Digitization is a social activity. It is not inherent in the stuff in the world and neither are the categories. Both the categories and the digitization are achieved by us.

The repair which occurs at the other end of the automation process is really digitization in reverse: When the unambiguous, concerted behaviour type of response of the machine is produced, we 'repair' it, converting its concerted behaviour output into something that looks like concerted action, which humans in the world relate to and respond to.

In effect, the machine is sandwiched between layers of human involvement in the process of automation. The human, by displaying charity to the artefact allows it to become an acceptable part of the world of regular human action, even though it only performs behaviour-specific action. These ideas assign great importance to the role of humans and far less to the machines. Humans are responsible for structuring action and digitizing the world. Behaviour-specific action, charity, digitization and repair are Collins's explanation of how machines that have not been socialised and do not participate in our form of life, become acceptable as pseudo members of society. Collins's view is based on the form of life and socialisation argument, and his ideas about skill and expertise are based on shared foundations and shared understandings (2). However, problems with the theory of behaviour-specific action have been identified: Slezak (1992) suggests that the theory is weak because Collins fails to specify in advance of any analysis what exactly in the world of human activity is behaviour-specific action and what is regular action. If the theory is to be used to look at what machines are doing, this distinction should be made at the outset. This is the only way of determining whether a computer has crossed the dividing line, mimicked a human engaged in regular action and jeopardised the theory. However, the distinction between the two kinds of action - those which are carried out on different occasions 'in the same way', and those which are carried out on different occasions 'in different ways' is not straightforward. The original division between the two may be seen as too static. Collins and Kusch (in preparation) are attempting to clarify the issues and the complexity of the terms involved:

They point out that variability between successive instances of an action occurs in *all* kinds of action. We do not carry out any action identically each time. In cases of regular action the variability can be great because part of the intention behind a regular act (say of writing a love letter or greeting a person) is to produce a variety of instantiations. This variability is part of what the action means to the actor. Whereas in cases of behaviour-specific action, the degree of variability is less, since variability is not part of the actors' intention. A degree of variability is tolerated, and the actors remain indifferent to this variability - it is not part of what the action means to the actor. As an example, Collins and Kusch describe turning a door handle. This is a behaviour-specific action, yet a variety of acceptable instantiations exist to which the actor is indifferent. Because of this indifference, actors doing behaviour-specific action could limit their instantiations to just one method that gets the job done. The degree of variability that is tolerated is different for each behaviour-specific act. For example, more tolerance is allowed when opening doors than when swinging a golf club. But some variations always occur.

Close analysis of this discussion reveals two new issues: First, behaviour-specific action can be mimicked by a machine that reproduces any one of the instantiations which fall within the margin of tolerance of the actor. Second, deciding - by watching - whether a behaviour represents either a behaviour-specific act or a regular act is not possible. To make this distinction, it is necessary to know the intention of the actor and whether they are aiming for a reproduction of an earlier action. This demands a knowledge of the culture or form of life of the actor; with such cultural knowledge, assumptions about the intentions of groups within a form of life can sometimes be made, but making assumptions about individuals' intentions is more difficult.

1.3ii The Medical Funnel

Blois' offers a slightly different explanation of the paradox of working computers. His medical funnel architecture is specifically designed to explain the kinds of tasks that automated medical devices can perform. Blois (1980) uses the funnel to represent the time from the first doctor-patient encounter to the formulation of a working diagnostic hypothesis (see chapter 2). The wide end of the funnel - point A - is the time of the first interaction, when the range of possible diagnoses is very broad, and the doctor uses a variety of clinical and wider world skill to approach the problem. Blois maintains that computers are not useful for this kind of task. Point B - the narrow end of the funnel - is reached when the problem has been narrowed down to specific details and where diagnosing the case requires the application of highly specialized medical knowledge, or a specific computation. At the narrow end of the funnel the problem has been reduced to a microworld format. No reference to the wider world, the context or to patient specific factors is necessary at this stage. Computers can be employed at this juncture because the problem has been sharpened to one that can be approached using a rule-based (behaviour-specific) strategy.

However, if computers are used at this stage, their output and their recommendations remain microworld-type recommendations. Until the output is assessed in the light of the particular patient under investigation, and analysed in the light of wider considerations, it makes little sense in the context of medical practice.

This is where I suggest an extension to Blois's theory: The funnel architecture needs to be modified to a 'venturi' shape. The bottom half of the structure then represents the 'unsimplifying' of the microworld-produced results into courses of action that make sense in the wider world. This new shape is shown in figure 4.1

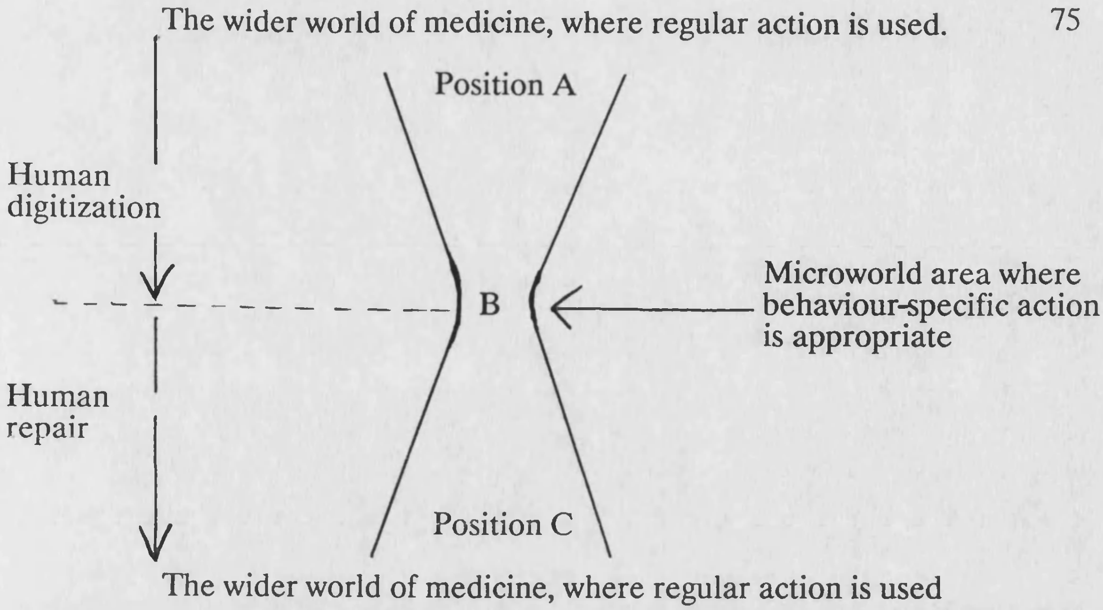


Figure 4.1 - The 'Venturi' architecture

Position C is reached when the decisions produced by the computer in the microworld of region B have been interpreted by a member of the medically qualified staff. At position C the influence of patient-specific and environmental influences have been introduced. Position C is where medical decisions join the wider world - it is like position A.

These two approaches to explaining what computers can do both emphasize the role of the human, and reflect a world influenced by human social activity. This fits in with the SSK promotion of the influence of social factors on all activity. On this basis, these two theories inform much of the analysis of the fieldwork which follows in section 2.

Is the selection of machines in medicine an appropriate arena for this investigation? There is confusion, both about the extent to which computers are used in medical environments, and the reasons behind the (apparently) low level of implementation:

1.4 Computers in Medicine - Opinion Divided

In the 1970s, Schwartz (1970) and de Dombal (1979) offered reasons why physicians need computerised help with their work. These reasons apply equally well today, and the field still thrives (3). Reports of the production of expert systems appear in the clinical decision making literature at the rate of one every month (Taylor, 1990). So what is the state of play in practice? Is the approach of SSK or of AI taking precedence? In the medical literature, opinions differ widely about the impending or existing impact of computers in practice. To summarise:

In 1979 Szolovits and Pauker claimed that "over the next two decades, computers will undoubtedly play an important role in the practice of medicine." (p. 1224). Maxmen's predictions of a 'post physician era' (see chapter 2) are even more optimistic.

However, other observers see things in a different light. Although medicine is a fertile field for research, the number of working systems does not suggest great success: Rector (1984) draws attention to this paradox, pointing out the rarity of systems in use. Ostberg (1988) puts his views more forcefully, stating that "There is obviously a serious gap between claims and reality of expert system applications" (p. 175). Sorgaard (1991) also discusses the disparity between the many apparently successful expert system prototypes and the low number of expert systems in actual use. Kingsland (1986) suggests that only three AI systems were in clinical use in 1986. Yet he makes the prediction that genuinely useful systems will be produced in the next decade.

Bearing in mind the research that is under way and the quantity of systems being developed (see the Index Medicus list), why are so few systems apparently failing to progress to the in-use stage? Are the problems of a philosophical and technological nature, tied up with the difficulties of assigning human knowledge to machines? Or are there more subtle social problems at the point of application of the new technologies? Is objection from medical staff in situ at the root of the apparent failure to install and run working systems?

1.4i A Human Problem?

Some writers suggest that the barrier to computer implementation is a human one: Taylor (1990) discusses new medical technology, stating that "the obstacles to its implementation in the real world of clinical practice are still human ones rather than the limits of technology." (p. 139). Maxmen (1987) sees the problem as one of user resistance, and Potthoff et al (1988) point out that "developers are more generally enthusiastic than prospective users and affected parties who ... only show a limited optimism" (p. 132). A more recent analysis by Anderson (1992) takes the same line, suggesting that the problem is a human one. He considered the attempt to install computerised diagnostics at the Royal Melbourne hospital. He argues that the social organisation of the hospital, the dominant culture of traditional diagnosticians and their control over the grammar of diagnosis made computerised diagnosis institutionally unacceptable. The clinical scientists and computer experts who advocated a change to computerisation could not secure it, because they were the weaker group in the culture.

1.4ii Or a Philosophical and Technological Problem?

Other writers have emphasized different types of problems, more philosophical or technological in nature: Sutton (1989a, 1989b) stresses that the difficulties facing machines are more to do with their lack of deep knowledge which underlies all surgical procedures. It is this lack of knowledge that prevents machines from performing at an acceptable level (p. 85). Similarly, Van Der Lei (1991) sees technological and philosophical problems facing designers who hope to enable machines to face unexpected circumstances in medicine. Pelosi and Lewis (1989) see problems for designers who believe that machines can intrude into the area of patient interviewing. Machines do not have the ability to undertake interviews in the clinical setting because they cannot be given the depth of understanding of the situation that is necessary - another philosophical problem. The system introduced by de Dombal at Leeds (1972) is reported to have achieved a 91% accuracy in initial trials (1972). But in 1989 when it was reassessed in a wider geographical setting it achieved only 42-57% accuracy. This also suggests that the technology is at fault - the machine cannot cope with a broader base of patients. It detracts from the argument that human user resistance is the only problem.

The philosophical and technological problems seem to hinge on the idea that "in much of medicine there is no consensus that defines proper therapy." (Van Der Lei, 1991, p. 1507). So it is difficult to achieve consensus about what to program. Although Engelhardt (1974) was not engaged in this particular debate, his statement that there are no pure and simple facts in medicine adds weight to these arguing from the philosophical, technological limits position. Whether the problems facing

AI in medicine are of this nature, or are more to do with human resistance is unclear. Forsythe (1992) offers an interesting conclusion: She says that users are often blamed for rejecting medical expert systems. But their rejection arises because the systems are built, designed and evaluated with scant attention to users' views or the nature of the setting in which the users work. (p. 95) Her analysis shifts the blame for rejection, from the users to the designers.

Beyond this another debatable point arises - where is the best place to use computers if and when they are suitably programmed?

1.4iii Where to Use Computers?

Opinions are divided about where in medicine computers would be best deployed. Some writers suggest that computers may be able to assist where "the problem is well bounded, which is computer talk to describe a problem for which large amounts of specialised knowledge may be needed, but not knowledge of the general world." (Nii, 1984, p. 111). A computer would not face unexpected circumstances if the problem were well bounded. Neither would it need to be equipped with knowledge of the wider world. By confining the computer to this kind of 'microworld' environment - the middle section of the venturi - the major mechanical problems are avoided.

Feigenbaum and McCorduck (1984) support this notion. They draw an analogy with apprentices in craftwork who are soon made aware that exceptions to rules are as numerous as the rules, and that what they have to do is get to grips with the spirit of the rule. Expert systems, they acknowledge, cannot yet manage this, and they are thus better suited to problems that are well bounded, as Nii suggested. In these cases, *applying* rules is more important than understanding the spirit of the rule. However, Feigenbaum and McCorduck also suggest that "Expert systems work particularly well when the thinking is mostly reasoning, not calculating - and that means most of the world's work." (p. 87) Can expert systems be best suited to problems that are well bounded and at the same time work particularly well at 'most of the world's work'? Does 'most of the world's work' occur in well bounded domains?

There is a further problem with Nii's suggestion: How many well bounded domains exist in medicine? We come back to the question 'where does medicine end and the real world begin?' If Blois is correct, by the time the problem has been reduced to one that is well bounded, requiring only medical knowledge, the human involved has already done a great deal of work sorting out the real world influences. Whether an expert system to finish off the job would be required or welcomed is then doubtful.

The alternative view is put forward by Madsen (1991), an opthalmologist. He suggests that there are problems involved in designing expert systems - such as deciding how much extra world knowledge to give the machines. In the light of such problems he believes that "The first patient care ESs [expert systems] to be used in optometry will be expert consultant systems for complex and/or unusual patient problems." (p. 120). The complex and unusual patient problems occur infrequently and when they do occur, the doctor may become engaged in solving a problem that she has not seen before. Solving this kind of problem requires a departure from run of the mill procedures, and the use of alternative strategies. This kind of problem solving does not take place within a microworld. Rather, it takes place at the boundary between medicine and the wider world. Capturing this kind of knowledge in an expert system forces the knowledge engineer to consider all the problems of representing wider world knowledge as well as more specific medical knowledge (4).

1.5 The Research Questions

Can machines be expected to intrude into all areas of medicine by small incremental steps? Or will the designers face problems? Most research in this area is theoretical - few researchers have considered the issue of computerisation in a real hospital setting from a practical point of view. An exception is Anderson's study (1992). His view was from a doctors/insiders standpoint. However, there is a major flaw in his study. By accepting that the computer was suitable to do the job it was designed for, he failed to look closely enough at how well the computer performed in practice. Anderson accepts the results of a clinical trial which suggested the computer achieved a 'success rate' of 69%, compared with physicians who achieved 42%. Anderson ignores the social influences at work in the interpretation of the trial result: In the trial, doctors were asked to make a diagnosis on information about patients they had neither seen nor examined. This is not what doctors usually do. Having accepted the study, Anderson was able to attribute the machine's failure to the 'power balance' within the hospital, and the possibility of the machine being unacceptable for other reasons was not investigated fully. Acceptability and competence of the machines has to be given more attention than was allowed by Anderson, and a grasp of the machine's performance in the practical setting is vital. If it does not do what the staff require and expect, it will not fit into the social group. Then, several alternative explanations to that offered by Anderson may emerge as reasons for the rejection of the new technology. Wyatt (1991) makes a similar mistake to Anderson. He suggests that "as with drugs, controlled clinical trials are the only way to assess their [decision aid's] impact on doctors and patients" (p. 1434). On the contrary - clinical trials are not the only means of assessing the impact of computers on staff and patients. Investigating the performance, acceptability and 'fit' of the systems in practice is another method. Yu et al (1979) recognise this point precisely. In their paper describing trials of the MYCIN system, they state that "The [trial] data demonstrate the program's reliability. However, further investigations in a clinical environment are warranted. Questions concerning the program's acceptability to practising physicians, its impact on patient care ... remain to be answered." (p. 1282)

Many questions remain unasked and unanswered. Deployment of medical systems has been less widespread than was initially forecast. However, there are enough systems in use to make this project necessary and feasible. The study reported here considers the role of staff dealing with computers and shows how well current machines fit into the social groups that they are a part of. The acceptability of the devices to the social group is of major importance in this research. Throughout the fieldwork the concepts of charity, digitization and repair are assessed as methods of fitting machines into the world of concerted action. The study considers the issue of where best to use medical computers. Are they best employed to perform behaviour-specific action or the 'bounded tasks' Nii suggests, within microworld situations? Do medical staff need help in microworld type domains? Or are machines more likely to take on difficult tasks and unusual medical cases that demand wider knowledge and reasoning? My adaptation of Blois's funnel is used to approach these questions. Can any behaviour-specific acts be identified, can any microworld tasks be isolated? How are microworlds created?

At present, no designers are attempting to make a mechanised general all-round physician - all systems are designed to operate within specialities. As such, all the systems are, to a greater or lesser degree, operating in a microworld within medicine. Is this in itself a feasible starting point? Specialised doctors (such as cardiologists, gastroenterologists or diabetologists, etc) recognise that consultation with colleagues from other specialities is often essential because very few patients fall squarely into one speciality. Are the designers making their first error by assuming that a system for, say dyspepsia (a condition falling in the gastroenterology speciality), can be developed without branching out into other areas of medicine? If there are problems associated with the introduction of machines, do they arise from philosophical or technical limitations, or are they more to do with human objections to new technology?

The research presented here was conceived and carried out in order to approach these questions from a practical perspective. The cases selected are machines in use in medical settings.

2. METHODOLOGY

The major factor influencing the methodological choice for this research was the necessity to look at developments from an insider's point of view. This is the best means of assessing the practical performance of the machines and their position regarding the social group and the social setting. Berger (1963) pointed out that "Sociology is not a practice, but an attempt to understand." (p. 15). I was concerned with gaining an understanding of the processes and actions occurring within the chosen field.

Debates about methodological choice, suitability and reliability are wide ranging but ultimately inconclusive. Here I will outline the rationale behind my choice and the advantages and disadvantages of that choice for this research.

2.1 Choice of Method

In 1984, Collins described participant comprehension as one type of participant observation. The essence of participant comprehension is maximum interaction with the members of the group under scrutiny. The aim is to become a 'native' member, overcome initial incompetences and gradually internalise the way of life of the group members. Socialisation into the form of life is necessary. Ideally, native competence is achieved as the researcher becomes accepted as a member of the group. The result is that the researcher's (or member's) own perceptions of the situation are as valid as those of any other member. The researcher is able to understand accounts offered by other members, and to formulate her own accounts.

Participant comprehension is the method that lies easiest with a view of human interaction in which perceptions, categorisations and rules are taken for granted within a social group. These things cannot be described, but are held in common by each individual within a form of life. Becoming a native of that form of life is the only means of accessing these taken for granted essentials.

There are difficulties associated with doing participant comprehension and gaining native competence. These difficulties can be to do with being a full time sociologist and finding the time to devote to pursuing native competence in another field. Other difficulties are to do with having competences that are acceptable within a group of sociologists, but untypical of the group under scrutiny. Beyond this, sociologists are outsiders unless they choose to investigate another group of sociologists. It is also difficult to know if or when native competence has been achieved. Finally, transferring perceptions gained as a native into terms acceptable to non-natives and other sociologists is a major hurdle.

Collins positions ideal participant comprehension at the end of a continuum of participant observation. At the other end of the continuum is unobtrusive observation. Any deviation from ideal participant comprehension will place the researcher somewhere else on the continuum. Any position is justifiable, provided the choices and compromises made are explained.

2.2 Participant Comprehension in this research

Gaining native competence in medical settings involves two processes: Gaining a general understanding of the medical form of life. By this, I mean knowing how to behave, dress and speak within any medical setting. This includes understanding procedures such as sterility in operating theatres, what to touch and where to stand. It also means grasping the hierarchical nature of staff grades, and taking on board the various degrees of deference shown to different members of staff. This type of competence allows the sociologist to blend into the hospital setting, and appear to know exactly what she is doing.

The other type of competence depends on the specific department in which the research is being conducted. If a cardiology department is the host, a knowledge of basic cardiac procedures and terminology will assist in achieving competence. Similarly, in a diabetic clinic a knowledge of blood sugar levels is needed, and for investigations in a gastroenterology department it is advisable to know an endoscope from a colonoscope. These more specific competences are less essential than the general competences outlined above.

I was especially well placed regarding native competence: I have wide experience of working in many NHS hospitals and a private clinic. I am more at ease in NHS settings (eg, see Hartland, 1990a), and undertook all the British case studies in NHS hospitals. My experience provided ready made native competence at the general level. I attempted to acquire specific competences relevant to each department as I proceeded through the empirical work. The extent of these specific native competences varied with each case study. It was greatest in the studies of the interpretive electrocardiograph machine and the automatic blood pressure measuring machine. It was rapidly gained in the gastroenterology department - this was largely due to the patience of the nurses and the senior registrar in the department who devoted much of his time to assisting me. Long explanations were also provided by a close friend, also a senior registrar in gastroenterology. Gaining specific competences in the area of jaundice diagnosis was more difficult, and in Utah, time constraints were the major problem. But in all the studies, my general level of familiarity with the medical world was invaluable.

2.3 Compromises Made

The ideals of participant comprehension were fulfilled in many respects. Most of the compromises I made were to do with time constraints, as it can be impossible to achieve native competence in a short time - it often occurs towards the end of the research (Halfpenny, 1979). When I could not achieve full participant comprehension and native competence, extended taped conversations were undertaken. These offer an approximation of native understanding, provided background research and investigations are completed beforehand. Adequate preparation is essential if the researcher is to present herself as an interested and knowledgeable contributor to the interactions. This is the only way to gain useful insights. It becomes clear when native understandings and frameworks have been internalised by the researcher, as she becomes able to distinguish between truth and lies on the one hand and jokes on the other. This is the case in ideal participant comprehension and during taped conversations.

The logistics of the fieldwork varied slightly in each case. This is shown in the fieldwork reports in section 2. In each case I undertook as much participation as possible, and supplemented this with conversations and interviews, observation and discussion.

Converting my native understandings into accounts acceptable to sociologists and lay people has been attempted by using colourful illustrations and notable quotations gathered during the fieldwork, to make my points. My examples have not been chosen randomly, but in a way best suited to conveying the understandings I gained within the medical form of life. To critics who feel that presenting selected quotations constitutes a bias in the report, I make the following suggestion: My findings are replicable by anyone with the time and motivation to undertake similar participatory studies. They must, however, be prepared to develop the same native competences as I did, before making any comparisons with my results.

The next section of the thesis presents the five major case studies. My writing has involved some retrospective rearrangement of events so that each study is presented in the same way. The structure of these empirical chapters will be as follows: A brief explanation of medical terms and procedures relevant to each study is presented in order to familiarises the reader and clarify the background issues. A description of the machine and its purpose is then given. Following this, the empirical findings are presented in relation to the issues for investigation and the unasked questions outlined above. To summarise, these issues are: How well do various machines 'fit' into the medical social setting? What is the role of the skilled human workers, and how relevant are the notions of digitization and repair? Where best does each machine fit into medicine, where in Blois's funnel does it work best? Is it only reproducing behaviour-specific action? Is the machine hindered because it is not designed to cross speciality boundaries? These preliminary observations will be drawn together in the summary chapter (chapter 10) where an attempt will be made to decide whether the barrier to the adoption of AI in medicine is a technological and philosophical one, or a human one.

The results of the study will add to the literature covering skills, automation of skills, deskilling, AI and SSK. The findings will fill the practical gap in the theoretical literature on AI in medicine. The optimum areas in which computers should be applied in medicine and the kinds of jobs they can do best will be discussed. The conclusions about the extent to which medical tasks can be computerised should inform the AI debate at a wider level.

FOOTNOTES TO CHAPTER 4

1. Slezak uses the example of cookbook recipes. These, he says are not fully explicit and neglect a myriad of explicit details. But, he says, "these [details] *could*, nevertheless, be made explicit." (p. 188).

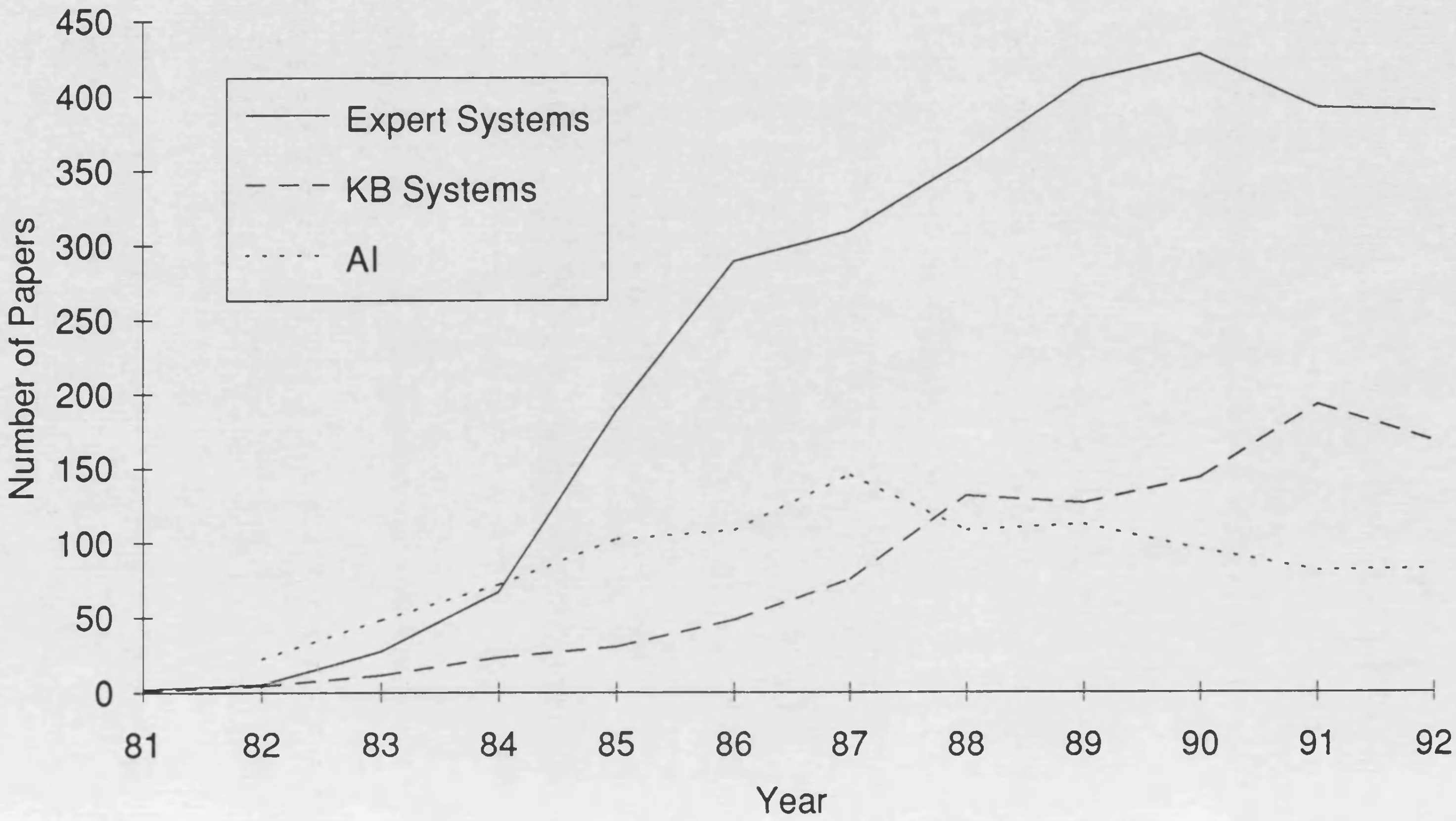
2. Collins clashes with Dreyfus on many of the issues here. Although both adhere to the later philosophy of Wittgenstein, the two authors hold different views about how socially constructed the world is. Dreyfus (1992) argues that if a domain has a structure that we can formalise, then a computer will be able to perform within that domain. If we do not know the structure of the domain, we will not be able to represent the structure in a computer, and computers will not succeed in that domain. The structure of a domain is a matter of how the world is; it is not something we influence. Collins disputes this. Furthermore, in their discussion of the acquisition of expertise (see chapter 3) Dreyfus and Dreyfus (1986) dismiss the notion of a computer performing at expert level because of the intuitive, holistic methods employed by people making decisions at this level. Collins suggests that the Dreyfus brothers' description of expertise ignores the social and collective nature of expertise which is gained through socialisation in a form of life.

3. The graph in figure 4.2 shows what has happened to the number of publications in the area of 'Good Old Fashioned AI', (which is defined here by Evans (1993) as expert systems and knowledge based systems). The number of papers increased steadily up until 1990 and then maintained a stable rate. Evans presents this as evidence of the high level of research activity in the areas of expert systems and knowledge based systems. The major area of application of these kinds of systems is the medical field, so it seems reasonable to suggest that the activity in the medical expert system field is still running high.

4. There are two problems involved in suggesting that expert systems will work best at complex, unusual problems: The first problem is that the relevant unusual problem has to be included in the database, which is less likely the more rare the case is. The second problem is that if unusual cases are included in the database they may be offered as diagnoses far too often. (See chapter 5 for a more detailed explanation of this point with reference to the GLADYS system).

Figure 4.2

So, what happened to GOFAI



SECTION 2

CHAPTER 5

THE GLASGOW DYSPEPSIA SYSTEM (GLADYS)

CHAPTER 5 - THE GLASGOW DYSPEPSIA SYSTEM (GLADYS)

1. BACKGROUND, PROCEDURES AND MEDICAL TERMS

This case study involves investigations into a computer designed to engage in a consultation with a dyspeptic patient, analyse the patient's symptoms and produce a list of possible diagnoses and a management strategy. Dyspepsia is "the general term for any symptoms concerning the alimentary tract, notably recurring or persistant pain or discomfort of the abdomen." (DTI publication, 1990, p. 8). The Concise Oxford dictionary defines dyspepsia as "indigestion". A more infomative definition can be found in the Dorlands Pocket Medical Dictionary which states that dyspepsia is "Impairment of the power or function of digestion; usually applied to epigastric discomfort after meals" (22 edition).

Dyspepsia is a very common complaint - between 4 and 5 % of all patients presenting at GP surgeries give dyspepsia as their reason for attendance. Of these 3 million patients, 10% - or 300,000 patients - are referred to the hospital Gastroenterology (GI) services for diagnosis and further investigation. Of those referred, 40-50% are eventually found to have a non-organic basis to their dyspepsia (Dunwoodie, 1987). One of the reasons for these unnecessary referrals is the danger associated with under-diagnosis: Since the underlying cause of the symptoms can be straightforward, or in some cases can be serious ulcers or carcinomas (Crean et al 1982), GPs often take a 'better safe than sorry' approach, in order not to miss any serious problems. It has been suggested that "diagnosing the patient that presents with dyspesptic symptoms can be difficult, time-consuming and expensive. The causes of dyspepsia are many and may overlap, definitive signs are few, and the path to accurate diagnosis is beset with potential pitfalls." (Smith Kline and French laboratories publication, 1988, p. 1). It is not a condition dismissed lightly by GPs.

In hospital GI clinics invasive tests such as endoscopy and barium meal followed by an X ray are used to determine the cause of the symptoms. But prior to these tests doctors should embark on a vigorous history taking that covers the occurrence and severity of the patient's symptoms. This is essential because the one characteristic that unites all dyspepsia cases is that the symptomatic history is of great importance (Knill-Jones, Dunwoodie and Crean, 1985, p. 204). Knill-Jones stresses this point, stating that "Symptoms are most important in the diagnosis and management of patients with Gastro intestinal disorders, in particular for those who present with symptoms falling under the broad heading of "dyspepsia."" (1986, p. 216). The best source of symptomatic information is the patients themselves. But the easy availability of test facilities can lead physicians away from the tedious and time-consuming task of accurate history taking and symptom analysis. Requesting a variety of tests is much quicker and easier than extracting a history and it has been suggested that this route is often ignored - that doctors do not appear to fully appreciate the value of carefully questioning their patients.

2. THE COMPUTERISED MACHINE

An objective in the era of cost cutting is to reduce unnecessary referrals made by GPs and thus reduce the number of unnecessary tests subsequently performed by hospital physicians. More rational theraputive, investigative and referral procedures are required (Knill-Jones, 1987), and one method of achieving this is by ensuring systematic and thorough initial history taking and symptom analysis at GP level. A computerised system for patient interrogation, 'GLADYS' - the Glasgow system for Dyspepsia - has been designed, which extracts a symptomatic history directly from a dyspeptic patient. GLADYS gathers the symptoms and various pieces of other patient-specific data through a series of questions which the patient answers using a specially designed keyboard (figure 5.1). At the end of the interrogation GLADYS produces a paper print out that lists the most likely diagnoses and decisions about the most appropriate management. It is, then, designed to perform two functions previously undertaken by doctors - history taking and diagnosis/treatment decisions. GLADYS has evolved over 20 years under the direction of physicians at the Glasgow Southern General Hospital and the University of Glasgow. Smith Kline and French laboratories have supported the project since 1981.

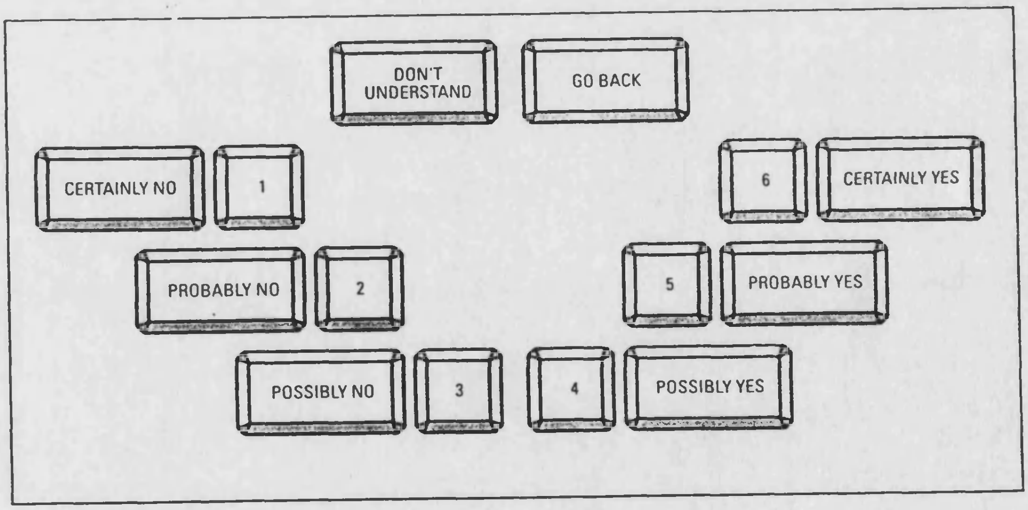


Figure 5.1

2.1 The Design of The Machine.

The GLADYS data base was compiled by a process of review of existing diagnosed patients. The information collected on each patient included interview and examination findings, endoscopy results, symptoms and final diagnoses. As many as 400 pieces of information on each of 1200 patients were included in the original database. This is still being expanded. The information was used to calculate the weight of evidence that a particular symptom or finding contributes to the presence or absence of a disease. Before the weighting system was perfected, several problems had to be overcome: First the patient population had to be identified using an exact definition of the disease. The definition of dyspepsia used was "Episodic, persistant or recurrent abdominal pain or discomfort, or any other symptom referrable to the alimentary tract except rectal bleeding and jaundice as the main symptom." (Knill-Jones, 1987). An explicit definition of each symptom also had to be used when compiling the data set. These two processes were designed to ensure that the area was well-defined and the data collected in a reasonably reproducible manner. For each case used, the final diagnosis had to be ascertained. This is an especially difficult task - some patients have more than one disease, and the level of certainty about the final diagnosis(es) can be possible, probable or certain.

Finally, the weights or scores were evaluated numerically and analysed according to a version of Bayes theory. Simple Bayesian analysis was not acceptable because it assumes independence of symptoms, which is not always the case in dysepsia. Instead, a process called 'logistic discrimination' was used with Bayes, which allows for symptoms that are dependent, and does not produce overly optimistic estimates (Knill-Jones, 1987). This mathematical approach was originally suggested by Card (1967) and in 1987 Dunwoodie stated that "although we seldom define and present evidence in this way, this is the path we follow in diagnosing" (p 162). McCartney (1987) also stressed the mathematical nature of the diagnostic process used by doctors, stating that it "bears a remarkably close resemblance to successive applications of Bayes theorem." (p. 1329).

The proponents of this 'weights of evidence' approach conclude that in areas where a deductive formula is insufficient but where extensive data is available, it is the most suitable approach. It has the advantage of overcoming many of the problems of purely statistical techniques whilst providing a valid probabilitstic output. In their analysis, Spiegelhalter and Knill-Jones (1984) suggest that dyspepsia is a restricted domain, where a large quantity of data is available, so it is a suitable area for the weights of evidence strategy.

Dyspepsia diagnosis was chosen for computerisation because it is an ill-defined area with overlapping causes, and because there are few definitive signs, and a wide range of diseases are possible (Dunwoodie 1987, Knill-Jones, Dunwoodie and Crean 1985). The DTI (1990), and Crean (1988), suggest that GLADYS'S greatest utility is achieved in difficult cases or cases where a number of diagnoses are possible.

GLADYS'S printout offers a list of diagnoses, attaching a level of probability to each of these. GLADYS also selects one of these diagnoses as the 'top management decision'. This does not necessarily correspond to the diagnosis with the highest probability, but the top management decision is the diagnosis that the computer considers to be in need of most urgent attention. A sample GLADYS print out is shown in figure 5.2 (p. 133/134).

It is claimed in the DTI document that a majority of patients prefer the machine to a normal consultation - they can spend 30 minutes with the computer as opposed to 6 or 7 minutes with a doctor. It is also suggested that patients are franker when dealing with the machine (1). Smith Kline and French (1988) state that patients respond well to GLADYS, and Dunwoodie (1987) specified that 82% had a favourable attitude to the machine, with 40% of patients preferring it to a doctor. The fieldwork will show how useful these figures are.

Four levels at which GLADYS can be described as 'in use' are set out by the DTI authors:

Level 1, as a means of gathering more information for the data base.
Level 2, being used in parallel with a specialist for development and improvement of design.
Level 3, in trial, which involves using it in operational form with an awareness that it is not fully proven.
Level 4, Used in its fully operational form without additional specialist assessment.

GLADYS is being used in Glasgow at levels 1, 2 and 3. Two other outpatient departments and a GPs surgery in Scotland are using it at level three. Beyond this, eight other UK hospitals are using it on a trial basis (level 3). One of these hospitals is in South Wales where the fieldwork for this study was conducted. GLADYS has been used in the outpatient Gastroenterology clinic at this hospital for about 18 months.

3. THE FIELDWORK

The fieldwork developed into three components:

1: Interaction with the patients using the computer, and with the computer (with myself acting as a phoney patient) was undertaken.

2: Contact was made with doctors and nurses involved with the clinic and the computer.

3: The third part of the fieldwork was involvement in the comparison of the computer's performance with that of the doctors. For this stage, a final doctor's diagnosis is required, and comparisons with the computer diagnoses are made. The idea is that mistakes and anomolies will be sorted out by the designers and then the symptomatic weights of evidence of a Welsh sample incorporated into the data base. This part of the fieldwork raises issues more to do with the mistakes the computer makes, and the reasons for these mistakes.

This three part strategy was adopted to see how well GLADYS functions in a practical setting.

My initial approaches to the consultant gastroenterologist at Cardiff explained my objectives, outlined my links with the Cardiff hospital and my acquaintance with the procedures involved and other staff known to both of us. Such an informal, 'back door' approach, mentioning previous experience in hospitals and other contacts, works well with medical professionals who are willing to by-pass bureaucratic access procedures if they feel that you are sufficiently familiar with the general medical form of life. I was invited to visit the clinic, talk to patients and staff, use the computer and generally take part in the clinic. My contact was to be the Senior Registrar on the gastroenterology firm. My level of specific competences and understandings of Gastroenterlogy increased as I began spending time in the department. I supplemented this with informal conversations with gastroenterologist friends and I undertook some basic reading.

3.1 Practical Use

GLADYS is switched on and set up by the nurses. They also explain the study and the procedure to the patients and generally oversee the smooth running of the computer in the clinic setting.

The computer requires basic information on each patient - age, hospital number, sex and marital status. The nurse explains how to operate the keyboard to input the 'yes', 'no' and 'don't understand' responses, and how to key in responses when the computer offers a range of alternatives. The nurse sits in on the first few questions, then normally leaves the clinic room.

My first encounter with GLADYS was in a mock consultation where I took the role of the patient. The questions are arranged in ten sections - introduction, primary symptoms, pain, bowel, heartburn, nausea, appetite, wind, weight and general. Top level questions are asked in each section and if answers are confirmatory, deeper level questions are asked. 350 questions are available, about a quarter of which are answered in each case. A reading age of 12-14 is assumed and the text is displayed at a controlled speed so as not to overwhelm the patient. My first reaction was that the questions were irritatingly long winded. Enquiries were made about family history - have your father, mother, uncle, aunt, brother, sister, son or daughter ever had ..... or .....? By the time the second condition was presented, I had forgotten the list of relatives I was supposed to consider. There were questions about when things occured - before breakfast or in the morning after breakfast - what if you don't eat breakfast, I wondered. Because I did not have a real complaint, my consultation became more and more confused as I gave incompatible and unsuitable responses. I then decided to look instead at real patients with real complaints using the machine.

The nurses send in a selection of patients to use the computer. Not all the patients were deemed 'suitable'. The staff explained that they chose those who appeared best able to deal with the machine - not too old and definitely not those appearing confused:

> "We try to pick suitable patients age wise. Its questions can be very confusing to patients. It usually takes half an hour to an hour, it really depends on the patient."

The consultant had suggested the same in his invitation letter to me, when he said that

> "the more intelligent ones deal with it in 10-15 minutes but some of the others need help and may take up to 40 minutes."

The nurses do not explain the difference between the possibly, probably, and certainly options for yes and no responses. This, it was feared, would confuse the average patient:

> "Possibly and probably, how do you define the difference between those two? We just find it totally confuses the patients and they don't know which one to press"

It also took more time, and as the nurse explained,

> "He [the consultant] likes four patients on here in a morning. I mean, we get all age groups in here, we don't only get youngsters".

So, responses are not differentiated into probably or possibly or certainly, and any of the three can be selected for yes and no. The patients are left to use GLADYS in the knowledge that they will have a 'proper' consultation with a physician afterwards.

During my observation of patients using GLADYS, two major problems were immediately apparent. First, the patients continually sought interpretation of the computer's questions, asking me to explain: For example, GLADYS'S first question concerns the patient's main symptom, asking them to select from a list of 6 possibilities. Typical responses to this request that were aimed at me were,

> "Well, more than one of these are the main symptom, what shall I do?", or,

> "The answer to this is a bit tricky, what do you think I should say?".

Most patients were anxious about misinterpreting the questions and turned to me for advice. When the screen showed a diagram and asked the patient to indicate which numbered position on the diagram their own pain corresponded to, one woman pointed to her middle and asked me where I thought it was on the diagram. It soon became clear why the nurse had left the room. The patients were looking for human intervention and assurance throughout the interrogation. One nurse explained that to achieve smooth running in the clinic

> "The best way is when you are sitting in there to help them. But that's not a nursing job."

Besides, they did not have sufficient staff on duty for this.

The second problem was that most patients wanted to expand their answers and felt frustrated at being confined to a 'yes' or 'no' response. One patient explained that if a doctor had asked him the question that the machine posed about the benefits of a particular diet, he would explain that

> "at the start it did good, but after that it didn't matter. But I can't get that across by saying yes or no."

GLADYS requires yes or no responses; these were often not the natural answers for the patient. They felt unable to interact with the computer and the one way communication left patients feeling that they had not conveyed sufficient accurate information to the machine:

> "It asks you questions, and it's all black and white, yes or no"

complained one man. I enquired whether he had conveyed all he felt to be relevant. He was vehement in his response:

> "Oh no, definitely not. I mean a human could get more out of you. You could come in here on an off day and get fed up pushing these buttons whereas a doctor would say why are you looking that way today, are you OK? A machine can't ask that."

Some patients complained that the machine failed to ask them what they felt to be important questions.

> "It just doesn't ask proper questions." (2)

Another woman was more distressed - she said

> "It doesn't ask you if you have had gallstones. I have had my gallbladder out, but it didn't ask me that, so this interview doesn't show what I am really like."

As the sessions progressed, the nurses allowed me to give the initial instructions to the patients. I decided to explain the relevance of the possibly, probably and certainly buttons to see if this would lessen frustration. The first patient appeared to understand my explanation and began the consultation. The machine asked if she ever had heartburn and she spoke, to herself, 'oh yes', and pressed the 'probably yes' key. When I asked her why she didn't use 'certainly yes', her reply was

> "Well I don't know if it is heartburn. It's never been diagnosed as heartburn, I just assumed it was, I've had Brufen for it. The doctor doesn't tell me why he's given me Brufen he just gives it to me."

She was sure that she had heartburn, but this had never been confirmed medically. As she was unable to explain this to GLADYS, she chose to answer in what she considered a more accurate form. Clearly the notion of 'probably' is ambiguous here.

A third, related problem came to light as I talked with the patients about GLADYS. Many of them felt that the questions 'drove' their responses - if a confirmatory response to a high level question was given, it was followed up with more questions on the same theme. When a woman was asked whether she had difficulty swallowing, she said,

> "Well, I don't have trouble swallowing but I suppose everyone gets food stuck and has problems sometimes, don't they?"

She answered yes. GLADYS duly followed up on this response with a deeper level question on the same topic. At this point the patient became concerned -

> "Oh no, now it is asking how long I have had this problem. I didn't think that I had a problem. I really have messed it up now, I'll be afraid to say yes again."

When asked if the pain was on the left, a patient said to me,

> "It is difficult, is my pain on the left? Or on the right? I think it is in the middle."

I suggested that if it wasn't on the left she should say no. The next question followed up on this, and the patient turned to me and said,

> "Are they asking about the same pain, because I get pains all over, every day"!

There was no way of responding with 'sometimes' or 'occasionally', which all patients expressed a desire to do. These qualifiers may have helped some patients give what they considered to be better answers, but many other patients would still not have been satisfied. They continually referred to the differences between this and a consultation with a real doctor, where they felt they could 'explain things fully'.

The use of words like 'often' and 'discomfort' in questions also caused confusion - "what exactly do they mean by this?" was a common enquiry. When asked about the time of onset of the problem many people are unsure, and unable to say whether it was six months, a year or a year and a half ago. One patient wanted to say that it was when Iraq invaded Kuwait, but this, unfortunately, was not an option. Similarly, the question 'do you belch more than a normal person?' caused concern:

> "How am I supposed to know how much a normal person belches?"

I was asked. Patients did not feel comfortable interpreting the questions themselves. My interpretation, opinion or advice was sought whenever these problems arose. The nurse later told me that the clinics I attended were the quietest from the nurses' point of view. She said

> "This isn't normally how it is run, because we haven't been bothered by the patients, they would normally have been coming out to us and asking all the time."

My presence relieved the nurses of the burden of dealing with incessant queries.

Dunwoodie's estimate that 40% of patients prefer it to a real doctor is not confirmed by my experience. It is unclear whether the patients that Dunwoodie is referring to were assisted by nurses when answering the questions. Patients are frustrated when asked to give yes or no answers when they would normally give a discursive verbal description. But are the patients' concerns unfounded? Is GLADYS able to make useful diagnoses and management suggestions even though the patients are dissatisfied with the consultation? Does GLADYS perform in a way that is comparable with, and acceptable to, doctors?

3.2 The Comparisons Between Doctors and GLADYS

The senior registrar was in charge of the evaluation of the system and the comparison of GLADYS'S diagnoses with those of the physicians. The doctor's initial diagnosis and final diagnosis, and the computer diagnosis were to be noted on an evaluation form which would eventually be returned to the developers (see figure 5.3, overleaf). My involvement in this comparison phase was intended to highlight the extent of the acceptability of the system's diagnoses and decisions, and provide an

Figure 5.3

FINAL DIAGNOSIS: GLADYS COMPUTER INTERVIEWS (code in boxes) [ ][ ][ ]

NAME:____________________ Patient ID No. [ ][ ][ ][ ][ ][ ]

INITIAL 1.__________ [ ][ ] [ ] (Accept: 1. Certain, 2. Probable, 3. Possible)

DIAGNOSIS 2.__________ [ ][ ] [ ] If 'Organic pain' i.e. Code 20,

3.__________ [ ][ ] [ ] use coding list to specify: [ ][ ]

DR RESP FOR DIAGNOSIS: [ ] (Code:1.1st Cons, 2.2nd Cons, 3.3rd Cons, 4.Other Consult, 5.Registrar, 6.Known)

INVESTIGATIONS: Barium meal / endoscopy codes: (Code '9' for 'Other' for all investigations)

Oesoph Ba Meal [ ][ ] Stomach Ba Meal [ ][ ] Doud Ba Meal [ ][ ] Certain? [ ] 1.yes 2.no

Endos [ ][ ] Endos [ ][ ] Endos [ ][ ] Certain? [ ]

Oesophagus	Stomach	Duodenum
01. Normal	01. Normal	01. Normal
02. H. hernia	02. Gastritis	02. Doudenitis
03. Reflux	03. Chr. Gastritis	03. Scar/deformity
04. Oesophagitis	04. Erosive gastritis	04. Ulcer
05. Stenosis	05. Ulcer	09. Other
06. Cancer	06. Cancer	
09. Other	07. Ulcer cancer	
	08. Ulcer scar 09. Other	

Ba enema [ ] Sigmoidoscopy [ ] Colonoscopy [ ] ERCP [ ]

1. Normal 2. Diverticulae
3. Chronic Inflam.B.D.
5. Polyps 6. Cancer
7. Stenosis
9. Other

1. Normal 2. Motor abnormal
3. Haemorrhoids
4. CIBD 5. Polyps
6. Cancer 9. Other

1. Normal
2. Abnormal

Ba follow thru [ ] Laparotomy [ ] Other [ ] Codes: 1.Normal, 2.Abnormal

Gall bladder [ ] Ultrasound [ ] Codes 1.Normal, 2.Stones, 3.Dysfunction, 9.Other

Lab: FBC [ ] ESR [ ] LFT [ ] Occ Blood [ ] Other [ ] Hist. avail? Y/N [ ] Campyl. Y/N [ ]

Lab test code: 1.Normal, 2.Abnormal

Period of obs. (mths) [ ][ ] Psych referral Y/N [ ] Psych features present Y/N [ ] Drugs & GI (1.NSAID 2.Other) [ ]

FINAL 1.__________ [ ][ ] [ ]

CLINICAL 2.__________ [ ][ ] [ ] If 'Organic pain' i.e. Code 20,

DIAGNOSIS 3.__________ [ ][ ] [ ] use coding list to specify: [ ][ ]

Signature for accepting diagnosis ____________________ [ ] (Code as above)

Basis for acceptance?: Code Y/N 1. Natural history [ ] 2. Compatible investigations [ ]

3. Resp. to treatment [ ] Diag. may be changed? [ ]

GLADYS OUTPUT

COMPUTER DIAGNOSIS

1.__________ [ ][ ] Probability [ ][ ] Satisfactory Computer Interview? [ ]

2.__________ [ ][ ] [ ][ ]

3.__________ [ ][ ] [ ][ ] (1.Yes, 2.No)

GLADYS TOP MANAGEMENT DECISION __________ [ ][ ] Probability [ ][ ]

GLADYS 04/89

opportunity for the doctor to explain differences between human and computer decisions (3).

I paired GLADYS'S print outs with the relevant set of patient notes, I entered patient identification data onto the evaluation forms and sifted through the notes (4). As the doctor looked for the final Gastroenterology diagnoses, we chatted about any differences between that and GLADYS'S decisions. He tried to explain why the doctors in question had made the decisions they had, what caused any differences between that and GLADYS'S decision, and why the doctor was right, or in some cases, wrong.

These sessions were relaxed (5) and from my point of view, very informative. The doctor was interested in all my questions, which perhaps offered him a new perspective and a different way of looking at the information we had before us. In many cases the computer's decision matched that of the doctor. However, in several of the thirty six cases the computer gave a different diagnosis or management decision to the physician. Six of these are described below:

3.2i - Case 1:

In this case, the initial decision made by the GP and the hospital doctor was that the patient suffered from biliary colic. Biliary colic signifies the presence of gallstones and the doctor explained,

> "She has got gallstones and somebody sensibly has taken her gallbladder out."

GLADYS'S list of diagnoses were that there was insufficient evidence (.7) followed by irritable bowel syndrome , or IBS (.12) and progressive alcohol problem (.11). Why did the computer fail to recognise this patient's problem as biliary colic?

Was it because the patient was young, and biliary colic is unusual in young people? The doctor dismissed this:

> "No, quite a lot of people in their twenties get it. It [the computer] has screwed up there."

He was surprised at the machine missing this diagnosis - reading her notes, nobody had had any difficulty with the diagnosis. The case was medically straightforward, and her history of one and a half years of episodic lower chest pain discomfort fitted the diagnosis of gallstones. GLADYS had placed gastric ulcer, oesophageal disease, obstructive bowel disease, duodenal ulcer, nervous dyspepsia, alcohol problem and IBS all above gallstones. Gallstones was so far down its list of suggestions that it had a probability value of 1313:1 against.

In an attempt to explain the error the doctor suggested that

> "She may just be one of these people who has trouble - there is always the possibility that she put in duff answers."

In fact there was nothing about this patient to indicate computer illiteracy or that she may be "completely naff", as the doctor put it. The willingness of the doctor to put the blame on the patient rather than on the computer is interesting (6).

Had GLADYS produced this decision in a GP's surgery and the GP had followed the advice given, the patient would have become one of the 40% of 'saved' referrals. The top management decision was 'insufficient evidence'. The patient would still be walking around, with both her pain and her gallstones.

3.2ii - Case 2:

Here the doctor's diagnosis was of oesophageal spasm. Apparently this is quite a rare syndrome. According to the literature (eg. Smith Kline and French 1988; Crean 1988; Dunwoodie 1987) it is in these unusual cases rather than the run of the mill cases that GLADYS will be of most use. In this case GLADYS'S list of diagnoses was, probable duodenal ulcer (.68) - this was also the top management decision - IBS (.60) and organic bowel disease (.16). The doctor pointed out that the idea of GLADYS being of most use in rare cases

> "assumes that the rare cases are in the data base. But usually you can spot something that is rare because it just doesn't ring true."

He went on to explain that oesophageal spasm

> "is something that we come across occasionally, and mostly you are making a personality decision. This guy winds his oesophagus up and gets pain."

So in this case the physician's diagnosis was based partly on personality assessment and partly on the feeling that the symptoms 'just didn't ring true'.

The decisions of the GP and the hospital doctor were the same -

> "Well he certainly hasn't got a DU [duodenal ulcer]",

said the senior registrar I was working with. This patient had been given an X ray which showed that nothing else was wrong, so he was reassured and sent away. This was "Fantastically cheap" as the doctor pointed out. Had GLADYS'S advice been followed by a GP at the initial consultation phase, the possibility of a duodenal ulcer would have been investigated via endoscopy. The DTI estimate that this would have cost aproximately £150 (DTI publication, 1990, p. 10). This would have been an unnecessary referral to the hospital.

3.2iii - Case 3:

In this case GLADYS'S top management decision was possible organic bowel disease (.84). The clinic consultant had made a definite diagnosis of depression. This he treated successfully in so far as at her next visit the patient

> "has got better, she has put on one and a half kilos. And she is no longer reduced to tears."

Why did the machine confuse depression with organic bowel disease? The doctor explained

> "The symptoms are very similar. A lot of depressed people have belly ache."

But why had the consultant involved been sure that it was not organic bowel disease?

> "First, because he has spent such a long time looking after patients and also he [the consultant] is good at recognising depression and treating it. He is more experienced with the treatment of depression, one of the few clinicians - Gastroenterologists - who are actually brave enough to do that - he makes a positive diagnosis that this is depression and he treats it."

If the diagnosis of organic bowel disease had been followed up the patient would have undergone various unnecessary tests, and her depression would have remained untreated. But the doctor's experience told him that this was not a physical problem and he acted accordingly.

3.2iv - Case 4:

The doctor's initial and final diagnosis were of definite hepatitis -

> "Everybody knows that this person has got hepatitis because that's what the referral letter says. What did GLADYS say ... It says organic bowel disease, ha ha ha."

This patient had a liver disorder and was not a typical dyspeptic patient. But as the doctor pointed out,

> "maybe having viral hepatitis gives you the symptoms that may well look like organic bowel disease - losing your appetite and feeling rotten."

How then, I asked, did he know that it was not organic bowel disease if the symptoms are similar to those shown in this case of hepatitis?

> "Because they are due to hepatitis"

> "But how do you know that?" I persisted.

> "Coz feeling sick and vomiting and going yellow are all part of it, but this [GLADYS] isn't designed to pick up on the going yellow."

This was the crux of the matter. The patient had a liver and jaundice problem and was from a domain external to dyspepsia. Part of the health professional's role is to choose suitable patients from the correct domain. The doctor or nurse should only assign to the computer patients suffering from dyspepsia. The doctor looked at it slightly differently:

> "Yes, but if I already know what is the matter with the patient how is the machine going to help me?"

3.2v - Case 5:

Here the doctor's decision was of recurring duodenal ulcer. GLADYS'S top management decision was of gastric ulcer (.24), although it suggested duodenal ulcer at .75. Apparently gastric ulcers and duodenal ulcers are difficult to distinguish between on clinical grounds and doctors often do not attempt it.

GLADYS recognised gastric ulcer as being more in need of urgent attention than duodenal ulcer, and recommended that the patient be referred for treatment of gastric ulcer. This would have involved various tests and an endoscopy at the hospital.

The doctor's management strategy was much different: on the basis of his experience of the patient, knowledge of the history and perception of the patient's habits, he chose an alternative course. The Senior Registrar explained:

> "The GP said this guy had a duodenal ulcer diagnosed by a barium meal ten years ago and every now and then his symptoms recur, does he need an endoscopy? Probably not because the symptoms recur every now and again and he smokes and he drinks ... this man is a plonker who drinks and smokes too much. He smokes fifteen fags a day and drinks aproximately a bottle of vodka a week, and in that situation sticking an endoscope down someone isn't really going to help."

He believed that this man should have been told

> "Look you have had this problem for ten years, you don't really look after yourself, you are actively making it worse. You have got to be kidding."

What eventually happened was that the man was managed clinically - no tests were done, and he was sent home.

In this case the diagnostic decision of the computer and the doctor was basically the same - ulcer, either gastric or duodenal. But the computer's advice to refer for tests for gastric ulcer was not appropriate in the light of the patient's habits and attitude. More important was to convince the patient that changing his lifestyle would improve his condition. Here it is evident that the computer may make the same diagnositic decision as the doctor, but this is just one stage in the overall treatment procedure. The management decision relies on the sensative reaction of the physician to other factors. The decisions made in the case of a 'well behaved patient' will be different from those made in the case of a patient who breaks the rules of acceptable conduct. The doctor's ability to make this kind of management decision is a basic medical skill.

3.2vi - Case 6:

The doctor's initial diagnosis was peptic ulcer. The patient was treated with drugs and when reviewed after ten weeks, was found to have normal endoscopy results. The patient had not returned to the hospital during the following sixteen months and the assumption was that the drugs had worked and the problem had been resolved. GLADYS'S suggestions were - as the doctor described it -

> "Stop being irritated and, [also] it can't decide."

He meant that the top management decision was 'counsel for nervous dyspepsia', and 'insufficient evidence' was top of its list of probabilities. This is an instance where the doctor decided on drugs and tests, whereas the computer advised counselling.

The reason for this difference was that the patient was a hospital employee. When dealing with members of staff the doctor felt that it was important to reassure them thoroughly - they are affected by the environment in which they work, and often feel that their symptoms signify the worst. The doctor felt that counselling would have been both inappropriate and insufficient - it would not have solved the medical problem and he did not believe that it would have settled the patient's mind. But was a desire to put the patient at ease the only reason that the drugs were given and the endoscopy done? Was there a clinical need to do that test? The doctor felt that there was:

> "I think he had quite convincing symptoms for pain that got better with treatment."

It is evident that the computer did not reach the same conclusion as the doctor. Whether the doctor made an accurate diagnosis of peptic ulcer is difficult to ascertain, since the drugs apparently cleared up the problem (if it existed) before the tests were done. However, this case shows once again that management decisions are based on cues that the computer has no access to - in this case, the fact that the patient was an employee and required complete reassurance that nothing serious was wrong.

When discussing the evaluation of the computer system, Knill-Jones, Dunwoodie and Crean (1985) who are closely involved with the GLADYS development, state that "the evaluation (therefore) has to be done in terms of [management] decisions made rather than diagnostic accuracy." (p209). In five of the six examples given here, the management decision made **and** the most probable diagnosis produced by the computer both differed from that given by the doctor. In the other example (case 5), the diagnosis was the same but GLADYS and the doctor chose different management options. Whether we choose to look at diagnoses made or management strategies, differences between the doctor and the computer are clearly evident.

4. DISCUSSION

Of the thirty six patients we looked at, there were ten cases where GLADYS and the doctor offered the same diagnosis and the same management decision. In these cases, GLADYS *was* able to reproduce diagnoses that counted as accurate, and choose management options that doctors also choose. So a computerised approach based on symptomatic evidence and probabilities can sometimes produce acceptable results. But there were sixteen cases where the diagnoses and the management decisions given by GLADYS both differed from those given by the doctor. There were also ten cases where the doctor's diagnosis appeared on GLADYS'S list but GLADYS'S management decision was different to that given by the doctor.

There are several ways of explaining these ten cases: First, the doctor's preferred diagnosis may have been given such a low probability by

GLADYS that it did not figure as a contender for GLADYS'S top management decision (as in case 1). Secondly, GLADYS may have offered a diagnosis at a high probability, which the doctor did not suggest at all. This computer diagnosis would then be much more likely to appear as GLADYS'S top management decision. Thirdly, the doctor's management decision may have been based on considerations aside from the actual diagnosis - the patient's lifestyle or personal circumstances (as in cases 2, 5 and 6). These are factors that the computer does not assess when making management suggestions.

These ten cases highlight two important points about the GLADYS system. First, if the doctor's diagnosis is included on GLADYS'S list, it does not necessarily mean that the computer deemed it a 'likely' diagnosis. Take example 1. GLADYS included the doctor's diagnosis of gallstones at a probability of 1313:1 against. Clearly, gallstones was not a serious contender for top management decision. But, as the doctor pointed out, because GLADYS "is programmed to try and work out probabilities, it will always come up with something." GLADYS may offer a long list of diagnoses, some with a very low probability. In this way all eventualities are covered - the longer GLADYS'S list the more likely it is that the doctor's diagnosis will appear on it somewhere. Compare this with a doctor dealing with a patient. The doctor does not produce a long list of possible diagnoses but is limited to one or two alternatives. GLADYS is not constrained by this limitation.

The second point is that doctors do not always make decisions about management in the same way that GLADYS does, even when both doctor and GLADYS are using the same diagnosis. Doctors take other factors besides the diagnosis into consideration - the patient's social and personal circumstances, their lifestyle and habits all influence the management decision. The result is that patients with the same diagnosis may be managed differently and patients with different diagnoses can be managed in the same way. It depends on the judgement of the doctor. This is echoed in an editorial in The Lancet (9 December 1989) which states that "Any practising clinician knows that different patients with the same disease differ in both the reaction to the disease and the response to various treatments. The good clinician learns to tailor his treatment to the individual patient, and, when the orthodox does not work, to use his knowledge of the ways of the body to design an unorthodox method that does." (p. 1371). GLADYS does not tailor its decisions to individual patients, and consequently often makes decisions that are at odds with those produced by doctors (7).

Judging by the comparisons we completed, it seems that GLADYS does not continously provide diagnoses or management decisions that physicians are happy to accept. In order to explain this, two issues need to be examined. First, the process of computerised consultation. What exactly is occurring? Is it flawed procedure, that leads to inappropriate decisions? Secondly, is GLADYS hindered in its decision making because its operations are confined to the limited domain of dyspepsia? The next secion discusses these possibilities in turn.

4.1 Problems with A Computerised Consultation?

My analsyis shows that the computer needs specific, or 'digitised' answers to specific questions. Patients' preferred answers were often more complex than the computer would allow. The computer works on a black and white, yes or no 'digitized' input. Real patients have to do a great deal of work to convert their own human responses into this format. The staff were also caught up in this process of providing the computer with an input it understood - they initially chose the patients, ensuring that they fitted into the relevant medical domain (8). Then they chose the most suitable of these patients for interaction with GLADYS - not too old, not too young, not illiterate, not confused, not computer phobic etc. In effect this amounts to selection of a 'digitised' patient in terms of suitability for GLADYS, in order to make things easier for the computer. The nurses assessed patients in terms of 'who will be able to answer the questions quickly and efficiently, without calling on a member of staff for help?' What was to be avoided was a necessity for a 'domain expert' (in this case a nurse) to intervene and fill in the gaps between the user's (ie, the patient's) knowledge and understanding and the system's knowledge base. Careful selection of patients was intended to ensure that the users' competence and the machine's knowledge base were as close as possible. (See Collins, 1990, p. 101).

Staff intervention is also necessary after GLADYS makes its recommendations. The cases detailed above show that repair of GLADYS'S diagnoses and management decisions in individual cases is necessary if maldiagnoses and haphazard referrals are to be avoided. Collins suggests that it is human 'charity' that helps to make up for machines' deficiencies. We act charitably (9) as we digitise their input and repair their output. In this case, humans act on the real world and convert it into a formal sub-set that GLADYS can cope with. Then the formal output of GLADYS is reconverted - by humans - into terms and courses of action that make sense and are acceptable in the everyday human world. GLADYS'S contribution to the consultation process is interfaced between layers of human contribution.

The concept of digitization - fitting the ambiguities in the world and in the patients' answers into specified categories - is similar to the narrowing that occurs in the top half of the medical funnel: As the doctor conducts the consultation and formulates ideas using judgement, the narrow end of the funnel is gradually approached. At this narrow end, a 'computation' is enough to complete the job. This metaphor resembles the narrowing down of responses into a form that GLADYS can use to complete the diagnosis task. The human work needed to do this moulding of input is essential if an artefact such as GLADYS is to blend into a social setting and if the gaps separating the world of the computer and the real world of medicine as to be bridged. But GLADYS does not blend flawlessly into the social scene. The mistakes and maldiagnoses occur because during the human/machine interaction, a reduction of the human world of dyspepsia, discomfort, symptoms and signs to a form that GLADYS can deal with is not always achieved.

4.2 A Confined World of Dyspepsia?

The Senior Registrar responsible for overseeing the evaluation of the computer in Cardiff offered some interesting observations about the machine, its methods, its area of application and its suitability. This doctor had little to do with the clinic use of GLADYS:

> "Oh no, it drives me spastic, some of those questions. The best ones are the husbands and wives together saying 'No you don't dear, oh yes I do.'"

He felt that GLADYS is not sophisticated enough to deal with the area of dyspepsia because dyspepsia covers such a large area, and there are

> "a lot of things that can be the matter with somebody that presents with dyspepsia".

The problem is that dyspepsia is a broad subject, whereas GLADYS is too limited. He said of GLADYS,

> "It is not suitable here. It's like asking the flight computer on a DC 10 to mow your lawn for you. It is not what it is set up to do. So, it may do as good a job as the flight computer on the DC 10 would do on the lawn. It may be OK [to use a computer] if it is a discrete subject. Like jaundice maybe. But belly ache has such a variety of causes." (10)

The system demands input within defined boundaries. It is intended for use with patients in a certain category. The boundary of this category is drawn around 'dyspeptic patients'. Whether this discrete area exists in practical medicine is unclear. The Senior Registrar certainly doubted whether dyspepsia is discrete enough an area to be suitable for computer application. This would explain why GLADYS'S management decisions sometimes differ from those of doctors working with the same diagnosis. The doctor is not operating within a microworld, as GLADYS is, and often takes other wider factors into consideration when making management decisions.

4.3 GLADYS'S Objectives Questioned

The fieldwork suggests that if GLADYS was used at GP level the number of cases referred and tests performed would alter. But how desirable is such a change in practice? The doctor at Cardiff spoke at length about the prime objective behind GLADYS, which is reduction of referal rates. He saw this objective as based on a simplistic idea of what a referal is for. He felt that in practice a GP may send a patient to the hospital for a variety of reasons. They may really need diagnostic assistance, they may need to be rid of a patient that is impossible to manage at their level - some patients have more respect for advice that comes from a 'specialist'. He estimated that in about half of the cases referred, the GPs know there is no serious medical problem with the patient.

> "The reason for referal isn't just this bland 'we want a diagnosis', it's 'we want a hand managing this smoking drinking bloke who won't stop smoking and drinking despite me telling him to. You try and tell him to stop.'"

Similarly, in other cases, a patient may come in one day and say 'I'm fed up of this doctor, I wanna go to hospital and see a specialist'. The Registrar explained that

> "It doesn't matter what the computer says about this man having no chance of having anything the matter. If he comes along [to the GP] and says 'I've got this continual pain, sort it out', most GPs with any sense will sort it out [refer it] because that's what the person has come along for."

The hospitals exist to offer a service to the GPs and it is not a case of the hospitals trying to reduce GP referrals.

His view of the purposes of a referral clashes with the traditional 'rational' view of much of medical decision making. Rational accounts stress the scientific basis of clinical judgement, and underplay other factors that influence decisions about diagnosis and other procedures, such as reasons for referral and discharge. The rational model assumes that medical practice proceeds according to laws and rules, in order to reach appropriate, 'correct' decisions. As Gordon (1988) points out, "The claim that medicine is scientific serves almost as a "covering law" symbolising a universal, "objective" truth and legitimising the authority of the medical profession." (p. 259). Atkinson (1977, 1988) describes how orthodox medicine is maintained and instilled in medical students. Their learning is guided and 'stage-managed'. The object of this is to subtly condition students into using recognised procedures and acceptable methods to produce recognised 'facts' and acceptable conclusions. The same applies to GP referrals - the rationale behind a referral is often viewed in a similar rationalist light and 'scientific' reasons for referrals are assumed. In theory, the acceptable reason that GPs refer patients to hospitals is because they need assistance with a diagnosis or because they cannot perform the necessary tests in the surgery. GLADYS has been developed on this assumption. The alternative reasons for referral - including those suggested above - are dismissed. In practice, deviations from these idealised procedures are evident. Gordon (1988) discusses the role of clinical judgement, intuition and informal criteria in decision making - it is a far cry from the rational road to truth. The Senior Registrar's comments above suggest that reasons for GP referrals also wander from the rational road. Many unorthodox reasons for referral are evident and external social influences affect the process. As Wright and Treacher point out (1982, p. 7) medicine is permeated by social forces.

The doctor also felt that GLADYS'S probabilities were unhelpful because, as a doctor, at some stage he *has* to make a diagnosis one way or the other. He explained that

> "Odds in a sense don't matter greatly. Even if there's only a 1:20 chance that you've got carcinoma of the stomach you'd like to know that you haven't. I see our role in diagnosing all percent chance. Carcinoma of the stomach is very unusual in 20 year olds but missing it is a major disaster area for all concerned. If someone says 'but my computer says that there's only a 1 in 100 chance' it's not a great consolation if you are the one."

GLADYS'S use of odds for and against certain diseases were, in his opinion, of little use in the real world of diagnosis. He used another theoretical example, of a computerised diagnosis for jaundice that suggested a 95% chance of cancer of the pancreas and a 5% chance of gallstones. In such a case he said,

> "It doesn't change what I am going to do because I have got to find out whether you have got gallstones or carcinoma".

His job was to decide what was an acceptably low level of risk. In many situations it is simply unacceptable not to know for sure, and further tests are essential. He pointed out that,

> "Yesterday we thought somebody might have two extremely rare diagnoses but both of them matter because they are extremely rare, they are very difficult to diagnose unless you think of them. Unless you think this person has got Addison's disease or Wilson's disease you certainly won't diagnose it. But you have got to make sure thay haven't because if you don't diagnose it it is fatal and if you do diagnose it, it is 100% treatable."

So the rare cases must be included in the data base but this may be problematc if the data base is small, as the rare cases may be suggested too often.

5. SUMMARY

Clearly there is some doubt over whether the objectives of GLADYS'S designers are acceptable objectives from the point of view of medical staff. This is interesting, but not central to the research. How GLADYS fits into the medical world, how staff respond to it and how GLADYS as a case study relates to the issues oulined in chapter 4 is more important.

I have shown that GLADYS is hindered because the information it gleans from the computerised consultation is incomplete and because it is confined to the specific world of dyspepsia. This means it does not have access to influences originating outside this field. By contrast, doctors tackling the problems GLADYS is designed to tackle, utilise their knowledge of the medical and social world outside the dyspepsia field, at the consultation stage, the diagnostic stages and the management stage.

GLADYS fits into the social setting as well as it does because of the effort made by the staff and patients. Human contributions are required in order to ease GLADYS into the social setting of the hospital. These human efforts narrow down, or digitize the world so that GLADYS only has to deal with specifics from that world. In Blois's terms, GLADYS performs at the narrow end of the funnel (venturi) in the microworld of dyspepsia, and produces output in terms acceptable at the narrow end. The six cases analysed in detail show that this output is often unacceptable in the real world (at the wide end of the venturi) and human repair, interpretation and 'widening' of the microworld type output, is necessary. The digitization, repair and adaptions made to the output in order to move it to area C in the venturi, show that human skill and judgement is essential when GLADYS is employed to undertake the job of dyspepsia diagnosis and management.

But still GLADYS is not an acceptable social prothesis - despite the efforts of the staff, it does not blend in well to the medical scene. GLADYS'S problems are technological problems, not human problems. The technological and philosophical difficulties of programming a computer to interview patients and produce consistently acceptable diagnoses and management decisions have been clearly shown.

Relating the findings of this study to the theory of behaviour-specific action requires an advance specification of what is behaviour-specific action and what is regular action in the area of dyspepsia diagnosis and management. Only then will it be apparent if a machine crosses this boundary and challenges the basis of the theory - that machines can only mimic people doing behaviour-specific action. In this case it would be necessary to show that GLADYS'S inconsistencies arise because it is doing the behaviour-specific action bits but failing on the regular action bits of this task. But separating these parts is difficult - our detailed case analyses were done in retrospect, using the patients' written notes, and explanations of the diagnoses that the senior registrar provided after the event. He could guess which cases had been straightforward, but could not be expected to say whether they had involved behaviour-specific action or regular action on the part of the physician involved. We can see that in some cases where signs and symptoms are clear cut and indisputable it would seem that behaviour-specific action is sufficient to lead to a diagnosis and management decision. As one doctor pointed out, "many cases are pretty straightforward." In many other cases things seem to be less clear cut, as stressed by Knill-Jones, Dunwoodie and by Crean. In the more complex cases, judgement, intuition, experience, context-related factors, patient-related factors and social circumstances are all taken into consideration. This fits the description of regular human action.

As far as the theory of behaviour-specific action is concerned, I have neither validated it, nor rendered it invalid through this study. It seems that the best method of investigating whether machines are consistantly performing tasks that humans perform using behaviour-specific action is to isolate an area where behaviour specific action is regularly used. The senior registrar had suggested that diagnosing jaundice was more suitable for computerisation than diagnosing dyspepsia because:

> "Jaundice is a more discrete, defined area and it uses laboratory tests quite a lot ... and the other thing is the diagnoses are all 'hard' diagnoses - gallstones, cancer of the pancreas, they are all proper diagnoses, for all clinicians."

This implies that Jaundice diagnosis is more suited to a behaviour-specific approach than is dyspepsia diagnosis, and that its 'hard' diagnoses are less influenced by factors from outside the microworld of jaundice.

The next case study was chosen with these considerations in mind. It involves a system known as SOLUBILE. This differs from GLADYS in several respects: SOLUBILE is designed to diagnose jaundiced patients, the intended user is the doctor rather than the patient, and the reasons that jaundice was chosen as an area are interestingly different to the reasons that dyspepsia was chosen as GLADYS' domain.

FOOTNOTES TO CHAPTER 4

1. This idea is also evident in Feigenbaum and McCorduck's book (1984). They state that "Studies in England show that many humans were much more comfortable (and candid) with an examination by a computer terminal than with a human physician." (p. 117)

2. On my first visit two interesting questions were posed: First, GLADYS asked a man if he had ever been pregnant, then asked a woman if she smoked a pipe. The computer asks the sex of the patient early on, and these subsequent enquiries suggest a design weakness.

3. My initial task as 'assistant' was to study the evaluation forms of the first dozen patients that had used the computer aproximately eighteen months earlier and retrieve these patients' notes from the medical records department. Then the final diagnoses could be ascertained. After this the doctor would be able to complete the evaluation form by entering this final diagnosis and the computer diagnosis. I was shown the workings of the medical records department by a clerical officer, and I collected as many of the sets of notes as were available. Some were unavailable if they had been dispatched to another clinic or another hospital in the area. This was a dusty, time-consuming job, involving step ladders, tracer cards and a supermarket trolley to transport the notes.

4. Medical notes appear complicated and disordered to the uninitiated - outpatient notes are at the front of the file, with GP letters, test requests and results. Any inpatient notes are at the back of the file, sometimes reading from back to front, and sometimes meeting the outpatient notes somewhere near the middle of the file. My familiarity with the system enabled me to sift through the pages and help locate the section detailing the diagnosis.

5. We undertook the task on several consecutive Mondays in the staff coffee room of the Gastroenteroloy department. By this time I was well known and accepted in the department. We managed to analyse thirty six cases, which leaves the doctor with somewhere in the region of two hundred and seventy more to complete.

6. I will discuss this point further in my summary.

7. Two other useful examples which illustrate this point arose during the fieldwork: The first was a patient diagnosed by the doctors as having nervous dyspepsia. The Senior Registrar said: "A whinger, lots of symptoms, not much the matter with you. A worrier. IBS or nervous dyspepsia. What else can you call it." He was sure of this diagnosis, but the management decision did not reflect this diagnosis - he decided to perform an endoscopy in order to reassure this 'worrier' that nothing was wrong with him. Similarly a patient with severe symptoms had undergone a barium meal test that was normal. The doctor had to decide whether to trust the result of this test - which is about 90% accurate - or perform more tests. The doctor was certain of the best action: "You have the difficulty of how much to believe the barium meal. In this case this guy here just drank too much. He's got belly ache and goes around heaving. He's got eight pints of Brains [Welsh beer] inside of him." So despite the symptoms, the doctor chose to believe the barium meal results, accept that this patient's illness was alcohol related, and not initiate treatment for oesophageal pain. In this case GLADYS recognised the alcohol

problem, yet the management decision it recommended was to treat oesophageal disease. Clearly humans often prioritise things differently to the computer in terms of management strategy.

8. Case 4 as an example, shows the problems that arise when a non-dyspeptic patient is presented to GLADYS.

9. Two examples of a different kind of human 'charitability' toward the machine arose during my discussions with the doctor evaluating GLADYS. As described above in case 1, the computer misdiagnosed a young woman with gallstones. Rather than accept that the machine had made a basic error the doctor attempted to charitably explain away the mistake, suggesting - with no evidence - that the misdiagnosis may have been the patient's fault. She may have put 'duff answers' into the machine. Again with a patient who, according to his written medical clinic notes, drank just two pints of beer each week. GLADYS suggested that this patient may have a progressive alcohol problem. The doctor pondered over this and eventually suggested that "He may well have told the computer something different to what he told us or he may have pressed the wrong button and instead of how much do you drink a day, [as GLADYS asks] he has said how much he drinks a week." This degree of charity is especially surprising coming from this doctor who was less than whole-heartedly enthusiastic about GLADYS. The instinctive reaction seems to be to repair and charitably explain mistakes. It seems almost automatic. Doctors can generally detect 'duff answers' and recognise when a patient is mistakenly stating his weekly alcohol consumption rather than his daily consumption. A computer does not possess the necessary social skills for this. These are abilities that humans have by virute of their understanding of other people, their socialisation in the world and their immersion in the human form of life and the medical form of life.

10. This suggestion, that a machine will 'work better somewhere else', arose frequently during the fieldwork studies. Sceptics often felt the problem to be the choice of area of application, and did not appreciate that the problems may be more universal than this.

Figure 5.2

```
GLADYS              DATE OF COMPUTER INTERVIEW:    16-04-91
------
RECOMMENDED ACTIONS, DIAGNOSES, AND SYMPTOMS FOR PATIENT NO: 614625
-------------------------------------------------------------------


DECISION 1 Possibly Organic Bowel Disease : Refer to Hospital G.I. Service
--------------------------------------------------------------------------
DIAGNOSIS:   Organic Bowel Disease      P = 0.2900      ( 2  TO ONE AGAINST)

Top 10 Symptoms for the diagnosis       Top 10 Symptoms Against The Diagnosis
----------------------------------      -------------------------------------

(+52)Mucus                              (-83)Age < 40
(+45)No pellets                         (-33)History > 6 months
(+25)Does not suffer from nerves        (-25)Episodic pain
(+15)No pain exac. from bowel action    (-25)Constipation


DECISION 2 Possibly Irritable Bowel Syndrome : Give IBS Treatment
------------------------------------------------------------------
DIAGNOSIS:   Irritable Bowel Syndrome     P = 0.6300      ( 2  TO ONE ON)

NOTE THE FOLLOWING SYMPTONS & WEIGHTS ARE FOR BOWEL CLASS -

Top 10 Symotoms for the diagnosis       Top 10 Symptoms Against The Diagnosis
----------------------------------      -------------------------------------

(+207)Diarrhoea + pain                  (-138)Diarrhoea pain not main pain
(+139)Lower abdominal pain              (-74)Antacid relieve pain now or in p
(+108)High fibre diet successful        (-45)Main symp NOT Pain/Bowel
(+74)Pain relief from bowel action      (-38)No pellets
(+71)Constipation                       (-35)No flatus
(+44)Mucus                              (-20)No nocturnal diarrhoea
(+36)Borborygmi 'excessive'
(+35)Female
(+19)No night pain with relief
(+18)Pointing sign - hand used
```

Figure 5.2 (continued)

PATIENT'S IDENTIFICATION: 614625 DATE OF COMPUTER INTERVIEW: 16-04-91

****

DIAGNOSTIC STATEMENTS APPROPRIATE	SUGGESTED ACTIONS (If not already taken)
3 Possibly Organic Bowel Disease	: Refer to Hospital G.I. Service
11 Possibly Irritable Bowel Syndrome	: Give IBS Treatment

****

DISEASE	P	ODDS
Irritable Bowel Syndrome	0.6347	2 TO ONE ON
Organic Bowel Disease	0.2880	2 TO ONE AGAINST
Insufficient Evidence	0.2194	4 TO ONE AGAINST
Nervous Dyspepsia	0.0980	9 TO ONE AGAINST
Duodenal Ulcer Disease	0.0390	25 TO ONE AGAINST
Progressive Alcohol Problem	0.0145	68 TO ONE AGAINST
Gastric Carcinoma	<.01	171 TO ONE AGAINST
Gastric Ulcer Disease	<.01	270 TO ONE AGAINST
Simple Oesophageal Disease	<.01	533 TO ONE AGAINST
Cholelithiasis	<.01	1043 TO ONE AGAINST
Severe Oesophageal Disease	<.01	2440 TO ONE AGAINST

*** THE IRRITABLE BOWEL SYNDROME MAY BE LOGICALLY IMPLIED, IBS SCORE = 6 ***

PATIENT NO: 614625 SYMPTOMS PRESENT

Food makes pain worse
Pain relief from vomiting
Early repletion after meals
No family history ulcer
No pain exac. from bowel action
Mucus
Episodic pain
Borborygmi 'excessive'
Pain relief from bowel action
High fibre diet successful
Lower abdominal pain
Pain relieved by food/milk
Self-induced vomiting
Night pain with relief absent/rare
No Tagamet prescribed
Vomiting in previous episodes
Belch relieves pain
Eat < 15 mins after vomitir
No night pain with relief
Does not suffer from nerves
No pellets
No daily pain
Antacids relieve pain now
Constipation
Diarrhoea + pain
Pointing sign - hand used
No attacks of pain
Antacids relieve now or in
Smokes < 25 /day
Pain more sev. & frequent
Single/Sep/Divorced

SECTION 2

CHAPTER 6

THE SOLUBILE SYSTEM FOR DIAGNOSING JAUNDICE

CHAPTER 6 - THE SOLUBILE SYSTEM FOR DIAGNOSING JAUNDICE

1. BACKGROUND, PROCEDURES AND MEDICAL TERMS

SOLUBILE is a computer system designed to make a diagnoses on jaundiced patients. Jaundice describes the yellowing of the skin, sclera and excretions that occurs when there is excess billirubin in the blood, and a build up of deposited bile pigments. Jaundice is recognised as a symptom of liver disease, as it is the liver which secretes bile and the pigment billirubin. Generally, a billirubin of more than 20μmol/litre of blood is taken as an indication that the patient is clinically jaundiced, the 'normal' range being 3-13μmol/litre. The distinct yellow colour of the skin and eyes is easy to see.

The Consultant Gastroenterologist responsible for devising the SOLUBILE program offered the following broad explanation of why he had chosen jaundice as the area of application:

> "Because I see that there are certain areas where doctors are bad at making diagnoses. And where it is obvious to me that computers would be good, and jaundice is one of the fields. Because in jaundice there is a lot of data a lot of which is quite 'hard'. So for instance a lot of the biochemical data is quite hard data. The serum billirubin is the serum billirubin is the serum billirubin and there is nothing that the doctor or patient can do to screw that up. I specifically chose jaundice because A: it is well encapsulated diagnosis-wise and B: it has a lot of mathematical and numerical data which is obviously ideal for the computer."

He elaborated on what he called the 'well encapsulated nature' of the field of jaundice diagnosis:

> "this [area] is relatively small and well encapsulated. We use only twenty two diagnoses here. In the whole of medicine there must be thousands of diagnoses. Twenty two we chose as being a sensible number. It is a reasonable number clinically to provide a range of options and it is not so few that it is clinically meaningless, and it is not so many that it is a nightmare."

He also felt that doctors could be 'educated' by a machine in this field. It allows them to learn why a given disease is deemed more likely than an alternative and why some tests are more rewarding in terms of the information they yield. The consultant felt that sometimes doctors forget to ask essential questions during consultations, and using the system should teach them to automatically ask the important questions:

> "It enhances the clinical skills because although we are taught in medical school to do all these things, when you come to enter things into the computer you realise that you have left out a whole load of things. It has an educational role and I think that it is all part of trying to give the junior doctors ownership of the process."

2. THE COMPUTERISED MACHINE

Unlike GLADYS, SOLUBILE was not designed for interviewing the patient directly. SOLUBILE gets its information from the doctor who has examined and interviewed the patient and performed the preliminary tests. The designer felt that

"Taking a history is a nightmare for computers",

and rather than have the system wrestle with consultations he designed it to manipulate specific pieces of discrete information - either numerical values or answers to specific questions (see table 1 overleaf). SOLUBILE presents a series of screens to the user, with requests for these forty seven different pieces of information about each patient. The doctor provides as many of these 'parameters' as possible, on the basis of the patient's history, examination and test results.

The computer diagnosis is based on the notion that diseases causing jaundice are characterised by the symptoms they normally produce. A probability for each disease being present can thus be calculated from an analysis of the symptoms and signs. The computer evaluates the information using Bayes theorem and produces a list of differential diagnoses, chosen from the twenty two alternatives that have been programmed into SOLUBILE.

The three goals of the designers of the program were to create a system that was

1. Easy to use, with a friendly interface and all input and output in plain language.
2. For it to produce useful recommendations about the most beneficial tests to perform, as well as a diagnosis.

TABLE I

PARAMETERS THAT CAN BE USED BY 'SOLUBILE' IN ITS DIFFERENTIAL DIAGNOSIS

History
Age of the patient (in years)?
Sex of the patient (Male/Female)?
Has the patient suffered from marked abdominal pain in this illness (Yes/No)
Duration of jaundice (in weeks)?
What is the duration of the patient's itching (weeks) (answer 0 if no itching)?
Weight loss (kg in last 3 months)?
Has the patient suffered appetite loss in this illness (Yes/No)?
Has the patient had pale stools during this illness (Yes/No)?
Has the patient had dark urine during this illness (Yes/No)?
Which category of drugs, if any, has the patient taken in the past 3 months (Type 'list' for the drug lists)?
What was the patient's alcohol usage in the last year (g per day/Guidance)?
Has the patient been jaundiced in the past (Yes/No)?
Has the patient come into contact with jaundice (Yes/No)?
Previous history of biliary surgery (Yes/No)?
Previous history of cancer (Yes/No)?
Previous history of biliary colic (Yes/No)?
Previous history of recent transfusions or intravenous drug abuse (Yes/No)?
Examination
What is the size of the patient's liver palpable below the ribs (in cm)?
Is the patient's spleen palpable (Yes/No)?
Does the patient have palmar erythema (Yes/No)?
Does the spider have spider naevi (Yes/No)?
Does the patient have Dupuytren's contractures (Yes/No)?
Does the patient have ascites (Yes/No)?
What was the patient's temperature (orally at admission, °C)?
Does the patient have encephalopathy (Yes/No)?
Does the patient have a peripheral neuropathy (Yes/No)?
Does the patient have signs of cerebellar disease (Yes/No)?
Does the patient have peripheral oedema (Yes/No)?
Investigations
Hb (g/dl)?
MCV (fl)?
White cell count?
Platelet count?
Reticulocytes (%)?
Prothrombin ratio?
Is any urine bilirubin present (Yes/No)?
Is any urine urobilirubin present (Yes/No)?
Bilirubin (μmol/l)?
AST (I.U./l)?
YGPT (I.U./l)?
Alkaline phosphatase (I.U./l)?
Albumen (g/l)?
Amylase (I.U./l)?
Does an ultrasound show dilated ducts (Yes/No)?
HBsAg + ve (Yes/No)?
Smooth muscle antibodies + ve (Yes/No)?
Antimitochondrial antibodies + ve (Yes/No)?
Alpha-foeto-protein > 100 I.U./l (Yes/No)?

3. For the system to be mathematically rigorous so that it requires minimal subjective user input, and remains accurate even with a small data base. (See Newman et al, 1988).

The first two objectives were aimed at making the system attractive and acceptable to users. The third objective arose from a desire to apply Bayes theorem more rigorously and thoroughly than previous programs had. (Newman et al, 1988, p. 183).

2.1 The Design Of The Machine

The original database was compiled from the case notes of 345 patients with a serum billirubin of greater than 20 μmol/litre, for whom a definite pathological diagnosis was available. As many of the 47 parameters that were available were entered in each case.

> "None of its data is based on consultant knowledge or textbook knowledge",

explained the consultant who masterminded the system.

When patient information is presented to the system it performs the mathematical analysis swiftly (details in Newman et al, 1988 p. 179-182): The designer explained that

> "It takes about four-and-a-half seconds on average to go through the maths. And it gives you the top five and their probabilities. And as a general rule what we have found is that if the probability is greater than 90% the chances of it being right are very high."

Details of all 47 parameters may not be available for every patient, and "SOLUBILE has been written to diagnose (and to learn) from whatever information is given to it, it is not necessary to enter values for every parameter." (Newman et al, 1988, p. 176). SOLUBILE also indicates which presently unavailable parameter or test result would increase the accuracy of the diagnosis, and to what new level of probability.

So as to help the user understand the rationale behind the diagnosis it is possible to ask the computer which parameters were most influential in the diagnostic choice. Comparisons between any two of the computers diagnoses, or between a computer diagnosis and a doctors alternative diagnosis can be made, to show why the computer considered one diagnosis more likely than another. Figure 6.1 (page 166) shows a SOLUBILE diagnosis sheet. The first section lists the five most likely diagnoses with probabilities. In this example, the second section is a comparison of the top two possibilities, alcoholic cirrhosis and pancreatic carcinoma. The third section lists those presently unavailable test results that would increase the level of certainty of the diagnosis.

The designer felt that this mathematical strategy was suitable, saying,

> "Jaundice is a straightforward, confined field, and what a doctor should do [to diagnose a case of jaundice] is to do what Bayes theory does, on specific bits of information."

The junior doctors on the gastroenterology firm (1) also seemed to accept that diagnosing jaundiced patients is a straightforward procedure, amenable to an algorithmic approach. One of the doctors explained:

"It is actually, a relatively mechanical thing. If you get this result then it is this type and when you get this result it is this type. It is quite a clear cut thing."

The most junior doctor on the team offered her opinion, saying

"I think, even as a houseofficer [HO], a junior member of staff, a jaundiced patient is a straightforward patient to deal with. I don't think there is any problem with diagnosing."

Another senior house officer (SHO) offered this interpretation of the processes involved in diagnosing a jaundiced patient:

"Somebody who is jaundiced, I mean, you walk into the room and it is bloody obvious what is wrong with them, and immediately you have narrowed down all the possibilities of what it could be and you figure it out."

The third SHO stated that

"Jaundice is quite straightforward. History, examination and a few blood tests can usually give you a pretty clear cut diagnosis. If doctors aren't very good at diagnosing it, it must be the way they are taught because it is quite a reasonably straightforward mechanical thing. We could almost sit down and write out the algorithm, 'how you arrive at the diagnosis in jaundice'. But you couldn't give the probabilities. I think that is much the reason why it [SOLUBILE] was brought out for the jaundiced patient - because it is actually a relatively a mechanical thing."

SOLUBILE is being set a task that is nearer point B than point A in the venturi model. The doctor has narrowed down the wider world - where jaundiced patients exist - to a narrower area characterised by specific responses and numbers. SOLUBILE is required to manipulate these discrete pieces of information and produce a diagnosis informed by the 345 cases in its data bank. According to the venturi logic, the machine should perform well at this (point B) level.

This may be the ideal type task for a computer - people can do the job using behaviour-specific action, and it lies near position B in the venturi.

The fieldwork was designed to assess the situation in practice: How acceptable is SOLUBILE to the medical staff? How well does it fit into the medical environment?

3. THE FIELDWORK

The research was conducted in a large general hospital in the Hampstead/Highgate area of London. My involvement with GLADYS had familiarised me with general concepts in gastroenterology, which overlaps with the area of liver disease, and I was at ease in my dealings with staff in the department. However, because of the travelling involved in this case study and because the computer was rarely used, full participant comprehension was not a viable option. Instead the fieldwork involved extended interviews and conversations with the designer and the relevant staff, and observation of the machine in use whenever possible.

3.1 Day to Day Use

Most 'intake' patients in this hospital are seen by a house officer (HO) or a senior house officer (SHO), and they were expected to use the computer. These doctors are notoriously busy and work long hours. A diagnostic device that saves them having to diagnose jaundiced patients could save valuable time and energy - a welcome addition to the available technology. The published statistics on the accuracy of the system are impressive (Alton et al, 1990, p. 52). In a 495 patient study SOLUBILE gave the correct diagnosis in first place for 78% of cases and for 92% of cases the correct diagnosis appeared in the top three computer choices. The study also compared the system with clinicians and the authors state that "SOLUBILE performed better, both in correct first place diagnosis and first three placing. No clinician reached the values achieved by the computer" (p. 52/53). The consultant encouraged his junior staff to take the patient history, do the examination, make a diagnosis and then use the computer diagnosis as a comparison, do the recommended further tests and use the computer again. In view of the accuracy figures I expected the staff to be pleased to use SOLUBILE, even though this was not compulsory. The Consultant designer explained that the situation in practice was much different:

> "No, no, no! They [the junior staff] have all sorts of other things to consider. This is another piece of work for them. Even if it only takes three minutes it is something else [for them to do]."

Apart from being too busy to use the computer he recognised that

"Some humans are terrified of computers. [That is] just how it is. Especially when you come to making a diagnosis. I think it is partly because the doctor likes the fun of making the diagnosis themselves or something."

Clearly he felt that the barriers to the implementation of the computer were human barriers. The medical staff were too busy or didn't want to use it.

The junior doctors admitted to using SOLUBILE infrequently and explained why: SOLUBILE is situated on the GI ward on the second floor of the hospital. Most patients arrive in casualty which is in a different area of the hospital:

"By the time they are got up to a ward or you have got time to go away and start using the machine you have got a fair idea of what is wrong with them anyway."

"It is stuck out of the way and you have to make a special point of coming up here."

"And it is in the endoscopy room. Half the time that is being used. They do endoscopies in there every morning."

So, few of the jaundiced admissions were entered into the machine because of this practical problem. All the junior staff had used it occasionally, and each described their experience and impression:

"In one case that I used it, the differential was between gallstones and carcinoma of the pancreas and it was 50.2 and 49, and the chances were that is was carcinoma of the pancreas, so it was no big surprise."

Here SOLUBILE produced a list of differential diagnosis on which two conditions were assigned almost identical probabilities. The doctor decided which was the appropriate diagnosis by utilising information about the patient's history, and dismissing the other diagnosis with the marginally higher probability.

Another doctor said that

"I thought it gave out completely zany results. I am sure that there must be some mistake in the algorithm. A young guy came in with jaundice and it said something which you never find in someone that young. I just put my own diagnosis in the patient's notes."

In this case SOLUBILE produced an unusual diagnosis which the doctor dealt with by ignoring it. Making a decision to ignore something is equally as important as making a decision to investigate something, and as much confidence and knowledge is required.

An incident that the third doctor recounted explained why she had not needed the machine to assist her in making a particular diagnosis. She said:

"I have had only one chap [with jaundice] but he had an enormously long history of severe alcohol abuse. And he was 35 so I mean it is pretty obvious what had caused the jaundice."

The doctor saw this as a straightforward case with an easy diagnosis of alcohol related liver disease.

These instances where the doctors had used SOLUBILE were the exception rather than the rule. SOLUBILE is not used often. What is the explanation for this?

3.2 Why SOLUBILE is Rarely Used

Judging by these examples and other comments made, it seems that there are three possible reasons why SOLUBILE is used so little by the staff it has been designed to help:

First is that the doctors do not need help in their diagnoses of jaundice. They see jaundice as easy to diagnose. So, they do not consider a computer to be useful or necessary. One of them felt that:

> "It [SOLUBILE] is looking at a very limited field. As the limitations [of the area of application] increase, the less useful it [the computer] is."

The doctors felt that the system was wasted in the jaundice field:

> "It would be better in a bigger area."

One of the junior staff was apologetic, saying:

> "I am really sorry about this, sorry that we can't say 'Hey it's great, we love it'. But it might be that if you could extend the program to

include other things, but a one-off program for jaundice, it doesn't actually, it is very limited. It would be better if you could use it for other things like rectal bleeding, or abdominal pain."

The staff continually stressed how much better it would be if the system had been designed for this kind of bigger area, or for the diagnosis of chest pain. In that area doctors admit to needing assistance, and the probability factor would be useful too:

"If you walk in and someone is 35 and they have got a bit of chest pain, and you are walking in and thinking right, is this something to do with their lungs or something to do with their heart, is it something to do with their indigestion. Are they hysterical, have they fallen over and cracked a rib? You are working on so many, a bigger field. Now something like that it would be more useful in."

The next doctor added to this, saying

"As a casualty officer, having done a casualty job, the number of people that you see going into casualty with chest pain and they vary from 17 to 117, with all sorts of varying history and stuff and at the end of the day, with the ones in the middle, the 35-40 year olds who do smoke but haven't got typical cardiac chest pain, it would be quite nice for you to sit down and put it all in and for it to come up with some probabilities and say, 'yes the probabilities of this patient having an MI [myocardial infarction] is 5%' or something."

A third SHO agreed, saying

> "If I went to see a young person with chest pain then I would be interested to know what the percentage probability is that she actually has ischemic heart disease."

Chest pain is a complex area of medicine. One of these doctors said of the field,

> "There is nothing you can really go on, you can do a blood test but it doesn't actually tell you anything."

That is why computer assistance would be appreciated. But they recognised the problems of assigning a computer to such a broad field:

> "I don't think it would work. You would have such an enormous amount of information that you would have to put into it."

A second possible reason why SOLUBILE is rarely used is that the probabilities which the machine produces are seen as intellectually interesting, but of little practical value when diagnosing jaundice. As one of the SHOs explained:

> "You have to be pragmatic about it. We are not interested in probabilities that are as weird and wonderful as 2 percent in the middle of the night. As long as you manage to see that if they need to go to theatre they go to theatre, and those that need antibiotics and resuscitation and those that don't, don't."

One of the doctors explained how medical staff make differential diagnoses in practice:

> "In a differential diagnosis you put the most likely first and it goes down to the least likely and if we really think it is out of the way we put in brackets, 'very unlikely'. But we don't actually quantify it because it is pretty meaningless."

It was generally felt that

> "It [SOLUBILE] does throw out some interesting possibilities that you haven't thought about and the probability of those things is so low anyway, like 0.9% or something, that you think, 'mmm that is interesting but not terribly relevant.' I haven't found it helpful. It just confirms my suspicions as to what the diagnosis is."

A third way of explaining why SOLUBILE is often by-passed, is that its diagnoses may be considered to be wrong by the doctors - as in the second case described above - and so it is regarded with suspicion. The research was not broad enough to determine whether this is regularly the case, and the trial results cited earlier in fact suggest that SOLUBILE gets it right more often than doctors. But as Lipscombe has shown (1990), clinical trials of expert systems often produce impressive results, whilst in practice the picture is different:

The figures in Alton et al's paper (1990) need to be examined closely. The first part of the trial was a simple 'reclassification' of 445 patients. This involves extracting a case that is already in the machine's database and requesting a diagnosis on it. For this part of the study the accuracy was given as 82% correctly diagnosed in first place and 96% diagnosed within the first three choices. This type of study is often regarded as meaningless - as the designer admitted, "If he [SOLUBILE] has already got the patient in the data base that is not really any good." It may also be argued that since the patient details are already in the database, SOLUBILE should get 100% of the reclassification diagnoses correct in first place. This type of test is really just a check on internal consistency.

In the next part of the trial, 'the prospective diagnosis trial', SOLUBILE showed slightly less favourable results. For this trial 50 cases were chosen "in an approximately similar disease distribution to that found in the data base." (Alton et al, 1990, p. 51). SOLUBILE'S first place diagnoses were correct in about 74% of cases and about 92% of cases were correctly diagnosed within the first three choices. But compared to the doctors' average accuracy, of 49.5% diagnosed correctly in first place, and 68.5% in the first three choices, the machine performed well. However, close questioning of one of the paper's authors revealed interesting unreported details:

> "They [the doctors] didn't have the patients in front of them, they had the same information as the computer, the sheets."

The doctors involved in the comparison were given the specific discrete pieces of information normally delivered to SOLUBILE. They were asked to make a diagnosis on the basis of information that another doctor had gathered, assimilated and interpreted. They were assigned the same part of the procedure that SOLUBILE is normally assigned, but were denied the opportunity to interact with the patient, interview them and examine them (2). SOLUBILE was in fact compared with clinicians who were asked to do something they would not normally do - make a diagnosis on the basis of another doctor's patient assessment. The designer dismissed the relevance of this point:

"My guess is that their diagnoses wouldn't have been any the better for seeing the patient."

It is difficult to ascertain how well SOLUBILE compares with doctors making diagnoses according to normal procedures, as this is not what the study detailed in Alton et al examined. The interactive part of diagnosis - talking to the patient and examining them does not only serve as a means of gathering clinical information. It is also a means of 'sizing up' the patient visually and mentally - assessing personality, state of mind and honesty. Such subjective analyses contribute to the doctor's overall diagnosis and management strategy. This part of the strategy is a finely tuned skill. Knowledge, confidence and practical ability are required if this initial part of the process is to yield the necessary information. The doctors who stressed how easy diagnosing jaundice is seem to have overlooked their contribution to this part of the process.

The 'hidden talents' involved in all stages of the doctor/patient interaction are essential aspects of the diagnostic process. These human parts of the process need careful examination if the value of SOLUBILE'S contribution is to be assessed.

3.3 The Human Contribution

When presented with a patient, the first task facing the doctor is deciding whether a patient is jaundiced. This is normally quite easy. As one junior doctor put it,

> "It is pretty 'either or'. And if it is borderline there is the blood test. But they [the patients] are usually duck's-foot yellow. Glowing in the corner."

But it is not always this simple. One doctor explained that some patients take weeks to present at a GP surgery because they have not noticed they are jaundiced:

> "Do you know, different patients, depending on their self awareness and their intellect, some patients are bright yellow and they haven't even noticed, and their wife has said, 'oh, you're getting yellow darling'. And some of them will come up at the first tinge. How do you get out of that. It is difficult is the answer."

Next, the doctor interviews the patient. The correct questions must be asked so that the necessary information is elicited. The doctor's personality, interactive skills, medical knowledge and experience all influence the quality and quantity of information gathered. The doctor examines the patient, and clarifies such things as the position and size of the liver, is the liver palpable? The doctor makes decisions about which blood tests and urine tests to request. Then the information is assessed in the light of the patient's history, social circumstances and personality.

The doctor is now in a position either to make a diagnosis or to answer the specific questions that the computer program poses, with specific digitized answers and numerical responses. All the information assimilated from the interview, examination and tests are collapsed into a form that the computer is designed to 'understand'. The system does not accept unusual values or responses that it is not programmed to deal with. Newman et al (1988) explain that "It is appreciated that some choices that SOLUBILE offers to its users may not be immediately understood by someone who is not used to SOLUBILE. For this reason SOLUBILE monitors all input made to it, and if a response is not reasonable it displays a help screen which should enable the correct response to be made." (p. 178). So SOLUBILE does not have to contend with unusual parameters that fall outside its range of comprehension. The system requires exact, digitized information and is entirely reliant on the doctor's skill and knowledge to provide that information. The idea is that digital information is free of subjective influences. But there are some problems with this - the designer pin-pointed questions that the program poses which are open to subjective interpretations. He said,

> "It starts off by asking some things that are quite certain, and then immediately you get into difficulty: 'marked abdominal pain', now, instantly we are into a difficulty about the interrelationship between the English language, human beings and computers. We have got a word in here - marked. Do you see that this is still open to uncertainty and subjectivity? So right at the start this is the crux of some of the difficulties of using computers."

The doctors' interpretation of what the patient says is essential. The doctor makes sense of the interview using prior knowledge of the patient, their social circumstances, their medical history. A great deal also depends on the doctors ability to relate to the patients personally, and understand their position in the same (human) form of life. So the input to the computer is subjective in so far as it relies on the interpretative skills of the doctor.

When presented with SOLUBILE'S analysis, the doctor may choose to perform the tests that SOLUBILE regards as most beneficial, or to ignore SOLUBILE'S recommendations. Similarly the doctor may take SOLUBILE'S diagnosis on board, or again choose to ignore it. Much depends on how sure the doctor is in her own mind about the diagnosis. The most important decision is whether to accept the possibility of a serious condition if it is given a low probability. Where does the doctor draw the line between an acceptable and an unacceptable risk? The doctor also has to decide whether to perform the suggested tests, taking into account the amount by which it increases the accuracy of the diagnosis:

> "Behind each of these [recommended tests] it gives a number of how much more it would increase the score. So it tells you that doing the platelet count is very worthwhile as it would increase it from 57% to 70%. So it sort of gives them an idea, and possibly for the cost of an MCV increasing it by only 9% isn't worth it. Or maybe it is?"(3)

Clearly, when SOLUBILE is used, the human involvement in diagnosing jaundice is still vast. In essence, the computer is being used only for a specific purpose - a subtask within the overall task of diagnosis of jaundice. It is used to manipulate specific pieces of data according to a mathematical procedure. This is seen as the best method of reaching the diagnostic decision.

> "What it [SOLUBILE] does is to take all the information and do what the human brain is less capable of doing and that is correlating all that information, to say, is a billirubin of 30 in combination with an albumen of 26, what is that? And we are very bad at doing that."

The designer believed that humans would make better diagnoses if, at this point, they too utilised a more stringent mathematical approach. This would be behaviour-specific mental action. Can doctors relate to this method of making decisions?

3.4 Behaviour-specific Action or Regular Action?

The designer thought that doctors found approaching tasks more methodically and mathematically, difficult. He suggested that was the reason why a machine would do the job more efficiently:

> "I don't think doctors think like that, which is a problem for doctors. What a doctor should do is to do some of what Bayes theory does, but the reality is that doctors don't do it. Which is perhaps why doctors aren't so good. For a start doctors are unable to store all the information in his brain and secondly because of that a doctor is unable to estimate the probability of any given diagnosis."

An SHO tried to explain how he thought the diagnosis process worked in practice:

> "I think we do all go around with Bayes theorem in our mind, but in an intuitive sense, you know, what is the most likely thing? And if you think that something is extremely unlikely then asking questions relative to that is not going to change that, and you are not going to believe the answers or you are going to interpret them in a different way."

Other doctors suggested that intuition played a much bigger part than mathematics:

> "a lot of it is intuitive. I was in casualty the other day and a lot of the time you walk into a cubicle and fix your eyes onto the patient and within about - I don't know - half a second, you have decided a lot of the time whether they need to come in or not, if they are that unwell. You fix your eyes on them and think 'oh, they're ill', or not, and so many things you make subconscious decisions I think, intuitively. Just by looking at them and seeing how they respond to you and that sort of thing."

Similarly, a senior registrar in Gastroenterology suggested that he could intuitively tell, as he walked up to a patient, or as they sat opposite him in a clinic, whether they really were ill, or "half way nuts". He claimed that if a patient wore tinted spectacles he would be inclined to initially class them as mad. On a more serious note, he said that just by looking at a patient, and "having a sniff", he could make "half a diagnosis."

The consultant agreed that a lot of diagnosis is based on hunch. Hunch and intuition are not behaviour-specific responses. Neither are the decisions doctors often make when armed with an understanding of patients' lives, problems, social circumstances and specific histories. In the cases where consideration of these unique factors influences the diagnosis made, regular mental action is used. Being able to make unorthodox decisions in the right cases is the mark of a good physician.

One doctor attempted to explain this:

> "a lot of our thinking is parallel rather than serial - it is not just one question at a time, it is all the input that has to be filtered all at once rather than that. Senior clinicians who go in, ask a few questions and say, well it is either this or this, and you think, how did they arrive at that, but it is just because they have seen so many things. And you find that when people get on but are still doing things systematically they obviously haven't got a clue what is going on."

Although the 'correct' diagnosis can be determined using an algorithmic approach, the best clinicians do not always do it this way. A knowledge of when to deviate from the behaviour-specific format is essential. SOLUBILE does not equip uninformed users or medical students with this knowledge, with 'gut reactions' or with intuitive acumen.

SOLUBILE uses a formula based mathematical method to do what people do using a combination of specific and ad-hoc methods. Humans do not only apply a formula, but rely on experience, hunch, intuition and patient specific cues; they respond to each patient in a unique fashion and use regular mental action. The designers of SOLUBILE are trying to replace this regular human action with a computer doing behaviour-specific action.

Because the task that SOLUBILE does is at the middle of the venturi, this should be an acceptable strategy. The narrowing down that the doctor has already done reduces the task to a microworld task, where behaviour-specific action should, theoretically, be an acceptable way of producing diagnoses on jaundiced patients. However, the fieldwork suggests that the results given by SOLUBILE were not always compatible with those of the doctors. Although the number of example patients involved in this study was small, discrepencies between the machine's output and the doctors', are clearly evident. Doctors sometimes disagree with the diagnosis given top priority, and sometimes choose diagnoses that SOLUBILE put in second place or lower. Explaining these discrepancies is not straightforward. But the crux of the issue is that SOLUBILE and the doctor are not really trying to do the same thing: In essence, SOLUBILE is designed to say which diagnosis is most probable given the available patient data. In contrast to this, the doctor is trying to make a diagnosis of what is wrong with this particular patient. The doctor may be interested in probabilities based on previous cases, but for each individual patient encountered, an individual diagnostic decision has to be made. One diagnosis may be mathematically more probable than another, but for each individual case the validity of that probable diagnosis is assessed, and a multitude of influences taken into account. Doctors' ultimate decisions are based on more than they can quantify in SOLUBILE'S categories. For example, one senior registrar suggested that

> "Doing the history and examination is the whole basis of clinical medicine, and with jaundice you have got some of the most difficult clinical signs. Like ascites. I would want to see for myself ... because by looking and seeing the patient yourself you decide yourself what's wrong or whether they are just half mad."

In a similar vein, a registrar said,

> "you focus all your questions and your examination to try to answer why they have gone yellow. So all the time you are asking the questions - there is no point in examining if you don't arrive at a diagnosis. You have to have a certain amount of clinical skill anyway and you have to develop some sort of relationship with the patient."

When deciding exactly <u>which</u> condition is apparent the doctor may engage in negotiations with colleagues or seek advice from more senior staff. Writing a list of probable diagnoses in the notes, with the actual diagnosis somewhere on that list is not acceptable in the real world of medicine. But proponents of machines like SOLUBILE would claim that producing the correct diagnosis in the top three possibilities is a major achievement for a computer. In practice, this is of little use to a clinician who has to decide where to draw the line, which is the most accurate diagnosis on SOLUBILE'S list? The doctor has to make a choice one way or the other and decide on the appropriate management strategy.

What the computer is designed to do - give the most likely diagnosis bearing in mind the patient information - is behaviour-specific action. What the doctor does - make a real diagnosis for a particular patient with particular idiosyncrasies - is not behaviour-specific action.

4. DISCUSSION

The designer and the intended users view the issues differently. The designer believes that doctors diagnoses would improve if they utilized an algorithmic approach. Doctors do not deny the validity of this approach, but they do not use this method once they have laid the foundations for a diagnosis - which is necessary whether they use SOLUBILE or not. Doctors isolate the problem, take a history, make an examination and take blood tests. Then they will formulate the diagnosis.

Staff were not suspicious of the machine, they simply saw little use for it, except as an interesting device that put numerical values on differential diagnoses. These numerical probabilities were of little practical use because doctors work on the basis that it is either this or that - a decision has to be made one way or the other. As one of the SHOs explained,

> "If somebody sees a patient and the differential diagnosis has fifty things, then people say, 'well obviously they don't know what they are talking about.'"

The doctor has to narrow down the diagnosis to specifics as soon as possible. It is the doctor's responsibility to interrogate, interpret and present the results to the machine. And then interpret the machine's results in terms of the particular patient under scrutiny. Just like with GLADYS, the machine is sandwiched between human layers which provide it with information in a ready-digitized format, and convert its digitized output into a form that is useful in the real world. The doctor deals with the output in various ways. As the examples above show, they choose between diagnoses that are given very close probability values, they decide to disregard spurious diagnoses, they chose the 'obvious' diagnosis in 'straightforward' cases. These decisions are only possible because of their experience and medical socialisation. They tend to play down their own role in digitizing the machine's input and analysing and repairing the output of the machine. Yet without the doctor on hand to assist SOLUBILE at the input and the output stages, the system would be of little practical use in a clinical environment. In order for SOLUBILE to blend into the clinical setting the contribution of a skilled human is essential.

5. SUMMARY

This case study has shown that the barriers to the implementation of SOLUBILE are human barriers. There are two ways of looking at why these human barriers exist. First, a great deal of regular action is required from doctors before they hand over to the computer. Making the diagnosis is the 'icing on the cake' rather than a tedious operation, and doctors are not keen to hand over this part of their craft to a machine (4). As the designer pointed out - sometimes doctors like the fun of making the diagnosis themselves. Investing time and effort in the earlier stages makes doctors keen to follow the task through to the end.

Secondly, the sub-task the machine does is not the area that doctors believe they need assistance in. Doctors may require assistance in making accurate overall diagnosis, but as explained above, that is not what SOLUBILE does. SOLUBILE does a behaviour-specific analysis of specific information, which in lots of individual cases is an insufficient basis for diagnosis. The real work of diagnosis, which is what doctors are interested in, is a sphere of action where patient-specific anomalies and unpredictable factors are important. Attempting to program a computer to perform at this wider level involves tackling insurmountable technical difficulties, because it is a job sitting at the wide end of the venturi. SOLUBILE is not designed to do this kind of task. SOLUBILE is only able to approach microworld tasks situated at the narrow end of the funnel. So part of the human barrier to use of SOLUBILE is based on an underlying technical shortfall.

However, the task that SOLUBILE is set is *not positioned near enough* to point B in the funnel. At point B, external factors have no bearing on the decisions made - they are encapsulated microworld decisions. But at the position SOLUBILE is operating, external factors still influence the decisions of human clinicians.

Designers of machines like SOLUBILE seem to be proceeding according to a fault-tree model of medical practice. Such a view ignores the patient-specific anomalies and unpredictable factors that doctors on the ground have to contend with, in a manner that utilizes artistic and intuitive elements, as well as scientific principles. The consultant involved here apparently adhered to this 'rational' view of medicine, and this view positions him at one end of a spectrum - the 'science' end. Practising doctors on the ward, who undertake the everyday work of clinical medicine often fall at the other end of the spectrum - the 'art' end. Gordon (1988) discusses these 'scientific' and 'art/clinical judgement' components of medical practice. She points out that "while science has long been the official knowledge of medicine, art and clinical expertise were long considered legitimate." (p. 282). She outlines moves intended to make medicine more scientific, and warns against such changes, suggesting that they constitute a threat to art and clinical judgement and as such threaten a valuable human resource (p. 284). Although SOLUBILE'S designer has attempted to use mathematics in place of clinical judgement, SOLUBILE does not threaten clinical judgement or art in this field, because it relies so heavily on the human contributions. SOLUBILE operates in a small area and human talents are required to reduce the real world into a form that SOLUBILE can deal with, and to translate the limited-world type output of SOLUBILE, into useful real-world information.

The stumbling block is that SOLUBILE attempts to use a behaviour-specific response to do what humans would do using a combination of behaviour-specific and regular mental action. As a result, its suggestions sometimes clash with those of human staff, and this makes them less inclined to use the device. If SOLUBILE is to produce acceptable results, it must be applied to tasks at the middle of the venturi - encapsulated microworld tasks where behaviour-specific action is sufficient.

This study was chosen because it seemed to fulfil my criteria of isolating a task that humans approach using behaviour-specific action. But closer analysis has shown that more than this is involved. The next case study was chosen with this consideration in mind: Blood Pressure measurement does appear to fit the behaviour-specific action model. Instructions for how to do the job are available - and as it can be reduced to rules an automated device should be able to take over the task. The next period of fieldwork involved investigations into the use of an automatic blood pressure measuring device in a hospital in Bath.

FOOTNOTES TO CHAPTER 6

1. 'Firm', or alternately 'team', is used to refer to all the doctors working under one consultant in one speciality. So, a full gastroenterology firm consists of the consultant, senior registrar, registrar, senior house officer(s), house officer(s) and any attached medical students.

2. The doctors were being asked to perform a job which, under these conditions, could only be achieved by applying rules and formulas. There was no scope for applying their tacit knowledge of the particular patient, because they had not interviewed or met the patient. They were forced into using a behaviour-specific approach. As Collins points out, people don't find this kind of action easy, and this test seems inherently biased against the doctors. Clinically testing medical systems is problematic. In this case the only suitable clinical test would be to compare doctors working without SOLUBILE with doctors working with SOLUBILE, and then to undertake a post-mortem examination and compare this result with the earlier tests.

3. An MCV test measures the mean corpuscular volume.

4. This is different to the case of factory robots that perform the sub-tasks of car assembly, or spell checkers that do a sub-task of article checking. These machines are used because those behaviour-specific sub-tasks are tedious. People are not sorry to see a machine doing these jobs. The part of diagnosis that SOLUBILE does is different. It is an interesting and exciting integral part of a longer process.

Figure 6.1

```
+------------------+              SOLUBILE              +--------+
|Hartland, Joanne.|              ********              |5678345|
+------------------+            16th Jul 1991           +--------+

                        +--------------------------+
                        |Diagnostic Probabilities|
                        +--------------------------+

        ( 9) Alcoholic cirrhosis                    Probability  71.8%
        ( 1) Pancreatic carcinoma                   Probability  17.7%
        (19) Acute alcoholic hepatitis              Probability   3.7%
        (20) Acute hepatitis (other)                Probability   3.2%
        (13) Hepatoma                               Probability   1.0%

+------------------+              SOLUBILE              +--------+
|Hartland, Joanne.|              ********              |5678345|
+------------------+            16th Jul 1991           +--------+

Alcoholic cirrhosis is 4.1 times more likely in this patient than
Pancreatic carcinoma.

The following are the main pointers to Alcoholic cirrhosis, followed by the
number of times more likely the parameter is with Alcoholic cirrhosis than with
Pancreatic carcinoma.

Patient is 27 years old                                      7.8

Patient has an albumen level of 26 g/l                       2.8

Alcoholic cirrhosis is more common                           2.5

Patient does not have abdominal pain                         1.3

+------------------+              SOLUBILE              +--------+
|Hartland, Joanne.|              ********              |5678345|
+------------------+            16th Jul 1991           +--------+

Based on what I know already, I think that the following questions are the ones
which are likely to be the most useful for improving diagnostic certainty.

 1)                              Platelet count (10^6/dl)?  (Score:  .7028)

 2) What was the patient's alcohol usage
                   in the last year (g per day/Guidance)?  (Score:  .6842)

 3) Has the patient had pale stools
                          during this illness (Yes/No)?  (Score:  .6715)

 4)                                            MCV (fl)?  (Score:  .6635)

The 'scores' given with the above questions are measures of how certain the
diagnosis is expected to be when the question is answered. A value of 1.000 is
perfect. For comparison, the current score is .5796
```

SECTION 2

CHAPTER 7

THE AUTOMATIC BLOOD PRESSURE MEASURING MACHINE

CHAPTER 7 - THE AUTOMATIC BLOOD PRESSURE MEASURING MACHINE

1. BACKGROUND, PROCEDURES AND MEDICAL TERMS

Blood pressure (BP) measurement is a common yet essential part of the daily medical routine. It is recorded on admission and regularly throughout a stay in hospital. This is the case for routine medical and surgical patients and for those in more critical conditions. On coronary care, intensive care and renal units the BP is recorded more regularly than on other units. Similarly, during surgery monitoring is intense, as fluctuations in BP can be an indicator of an imminent change in a patient's overall condition. The BP is expressed as one figure over another. The top figure (the systolic pressure) represents the pressure exerted when the ventricles of the heart are contracted. The bottom figure (the diastolic pressure) represents the pressure exerted when the ventricles are relaxed.

The BP can be measured either invasively or non-invasively. Invasive measuring is a complicated procedure: A fluid filled line is inserted into an available artery while the other end is connected to a pressure transducer and a VDU showing the pressure values. This procedure is used most often in surgery and intensive care units as it allows continuous monitoring, and because of the close proximity of the arterial line to the heart, gives results that are considered to represent closely the pressure inside the left ventricle. Non-invasive methods are used much more often. This familiar procedure is most often performed using an inflatable cuff wrapped around the upper arm, attached to a sphygmomanometer which measures pressure in millimeters of mercury (mm Hg). The cuff is pumped with air until the pressure it exerts interrupts blood supply to the arm. By slowly deflating the cuff and noting the pressure value on the sphygmomanometer when blood flow resumes - either by palpating (feeling) the radial artery or auscultating (listening to) the brachial artery through a stethoscope - the systolic pressure is ascertained. Further deflation to the point where the sound of the blood flow at the brachial pulse disappears determines the diastolic pressure (1). The BP is then expressed as systolic over diastolic. This "technique of listening to sounds is referred to as the auscultatory method." (Critikon publication II, p. 4) The method was developed around 1905 by Nicholai Korotkoff.

Normal values vary enormously depending on age, history, fitness and lifestyle. However, a very general guide-line often used is that an 'average' resting adult can expect a BP somewhere in the range of 120/80 mm Hg.

A nurse, a junior doctor, a more senior member of the medical 'firm', or an anaesthetist may normally be responsible for measuring the BP. Other paramedical staff - especially cardiac technicians and anaesthetic monitoring technicians - are responsible for monitoring the BP during various theatre procedures.

A working party of the British Hypertension Society has produced written guide-lines for taking the blood pressure in the correct fashion (see Petrie et al, 1986). The instructions for taking blood pressure can be set out - the following passage from Miller and Keane (1978) is, in effect, a set of rules for taking blood pressure:

> "Measurement of the Blood Pressure: The blood pressure is usually measured in the artery of the upper arm, with a

sphygmomanometer. This consists of a rubber cuff connected to a glass tube containing a column of mercury. Alongside the glass tube are numbers that indicate the height of the column of mercury in millimeters (25 mm. equals 1 inch). In some sphygmomanometers the mercury column is replaced by a gauge. The rubber cuff is wrapped about the patient's arm, and then air is pumped into the cuff by means of a rubber bulb. As the pressure inside the rubber cuff increases, the flow of blood through the artery is momentarily checked. The pressure within the cuff causes the mercury to rise or the gauge's needle to move.

A stethoscope is then placed over the artery at the elbow and the air pressure within the cuff is slowly released. The pressure begins to fall slowly. As soon as blood begins to flow through the artery again, tapping sounds can be heard through the stethoscope. This is the pulse. When the first tapping sound is heard, the systolic pressure is noted.

As the air pressure continues to escape from the cuff, the tapping sounds grow louder. A point is reached at which the sounds change suddenly to very soft and then disappear entirely. The point on the mercury column at which the sound disappears entirely is the diastolic pressure." (p. 130-132)

From this it seems that the 'same' sounds will be audible in every patient and the task is best approached using the same strategy every time. Humans carrying out this procedure could use behaviour-specific action. It follows - according to the theory of behaviour-specific action - that a machine should be able to do the job by following the same rules that a human follows when recording the blood pressure. A human recording the blood pressure and a machine recording the blood pressure would both be doing the same thing - trying to find out what the numerical value of the blood pressure is. This is not an attempt to make a 'diagnosis', but is a presentation of information that will assist the doctors making the diagnosis. Measuring the BP is a very small part of many complicated medical procedures, and does not appear to require reference to external factors, or the extensive use of regular action that those more complex procedures demand. It is positioned in the region of point B in the venturi.

2. THE AUTOMATIC MACHINE

Manufacturers have recognised and stressed the advantages of mechanising this job - a wide variety of automatic BP machines are on the market. The machine pin-pointed in this study, manufactured by 'Critikon', is the 'Dinamap 1846SX'.

This machine uses a procedure that looks similar to the hands-on, non-invasive technique used by humans - it inflates the BP cuff to a level that enables it quickly to determine the systolic pressure. Then it deflates the cuff in steps to detect the systolic pressure. The machine uses the same procedure in every case.

On the front of the machine is a colour-coded control panel. This allows the user to instruct the machine to take a single reading, serial readings at specific intervals from 1-90 minutes, or a rapid succession of readings over a five minute period. The range of acceptable readings is also programmed into the machine, and an audible alarm sounds when a reading falls outside this pre-determined range.

The job of the machine is to record the BP value and alert a professional if this value falls outside the programmed range. The Dinamap machine is presented as an accurate, fast and easy to use device (See Freisen and Lichtor, 1981). Critikon cite independent published papers that draw attention to the reliability and accuracy of the monitors. These papers generally conclude that "The instrument was found to give good results in a wide variety of clinical subjects and physiologic states" (Ramsey, 1979). The conclusion of Inoue's study (1988) is that the Dinamap monitor is an accurate means of measuring non-invasive BP in the operating room, and Epstein et al (1989) found the Dinamap monitor more reliable than the machine they compared it with, producing measurements that were closer to the auscultatory measurements on the chosen patients. It was also rated best automatic BP machine in a separate study comparing seven different devices (Health Devices Study, 1987). Grundy et al (1981) go so far as to suggest that "the Dinamap readings may in some instances be more accurate than intra arterial pressure measurements" (Grundy et al, 1981) (2).

According to the manufacturers, the Dinamap 1846SX has "set new standards in the measurement of Systolic, Diastolic and Mean Arterial Pressures and Heart Rate." (Critikon publication I). Critikon have also conducted their own studies comparing the Dinamap readings with invasive pressure recordings. The results show mean differences between the two methods of plus or minus 5 mm Hg. As Seaman (1985) points out, using a non-invasive machine is far simpler and safer than invasive methods and spares the patient the trauma of undergoing an invasive procedure.

The impression conveyed in the literature is that the machine performs well, providing quick, accurate measurements and effectively taking over the task from the human operative. The literature reports suggest that the machine is achieving high levels of acceptance and effectively doing the human task. The empirical fieldwork was designed to analyse the situation in practical settings.

3. THE FIELDWORK (3)

3.1 Practical Use

My first impression was that the machine performed well - on a unit where it was used regularly the nurse in charge reported that,

> "You put it on to automatic when you do it over a set period, and then you just leave it and it does it all."

It fitted in well enough to be left to do the job unsupervised. Another nurse, who had previously worked on an intensive care unit, explained how it had been used there:

> "There was no compunction in just setting it up and leaving it and trusting it to tell us ... I have found it very reliable."

This matches the sentiment expressed by Seaman in her 1985 paper. The title of this paper is 'should you trust automatic blood pressure monitors?'

Her answer was "Yes, they give consistently accurate readings, [and] save you valuable time."

The staff seemed satisfied with the machine and confident about the ability of the machine to do the job unassisted. Once set up, I watched the activities on the ward and saw that the staff routinely left the machine, only returning periodically to read the results off the paper print out, or to respond to an alarm. A member of staff on the geriatric ward summed this up, saying "This does the job."

The staff involved with the machine seemed to assume that taking the BP was straightforward, something that a machine can be allowed to do. To them, it was not a complicated task - any skills involved had become invisible to them.

But a closer analysis of the literature suggests that taking the BP is far more complex than this: Although BP measurement is a procedure utilized at every stage of patient care, and it is taken for granted that staff are familiar with the non-invasive method and the significance of the readings, there seem to be difficulties involved: Jamieson et al (1990) sum this up, saying "Measurement of blood pressure is one of the most commonly performed clinical procedures but it is beset with difficulties in terms of instrumentation, patient factors and observer differences." (p. 6435) O'Brien and O'Malley (1990) also suggest problems with the BP measurement technique, describing it as a "frail foundation" for managing and researching into hypertension. The problems associated with measuring BP, the factors that lead to false readings and guide-lines for measuring the BP are set out clearly in a document prepared by a working party of the British Hypertension Society (Petrie et al, 1986). In the

Critikon publication (II), twelve common sources of error associated with taking the BP are discussed (p. 7). These include faulty and badly maintained equipment, and variations in observers' techniques. It seems that a variety of methods are used by different staff, depending on where and when they were taught and by whom. This idea is echoed in Hatt's research (1992). Amongst her conclusions is the assertion that methods of measuring blood pressure vary widely. Techniques depend on the grade, gender and geographical location of the measurer, and are highly context specific. Differences between measurements made by the same member of staff on one patient, are also discussed. (See also Dawe 1993, Jamieson et al 1990, Short 1976).

A senior registrar I spoke to also referred to these differences. He called the variation 'observer bias', and explained what he meant:

> "If you ask a nurse to take a patient's blood pressure and ask another one two minutes later and another one two minutes later, they'd all make it different. It's notoriously unreliable."

He recognised that the methods adopted by different members of staff were variable, and he pointed out why this was important:

> "You take the blood pressure by feeling the radial artery and pumping the cuff up. That's the right way to do it. [But] Nurses listen [to the pulse inside the elbow, at the brachial artery] and when the sound disappears they call it the systolic. But the sound can disappear, and when you go on pumping it up it can come back in later on. It is called an oscillatory gap, which you miss. You don't miss that if you palpate the radial artery. And I have seen

> nurses underestimate the blood pressure by something like 100 mm Hg. It is possible ... and when do you say it is diastolic? When the sound muffles or when it disappears? There could be 5, 10, 15 mm of difference between that. If there is a significant difference you are supposed to record both. Some doctors don't. Nurses don't. Half of them don't even know whether they are doing it at the fourth phase or fifth phase."

It seems that the methods used to record blood pressure are context specific, patient specific and measurer specific. The activity is performed in a different way at different times, and fits into the category of regular action. According to the theory of behaviour-specific action, the task is thus beyond the capabilities of an automated device.

So what exactly is involved in measuring the blood pressure? Is it a situated action, as the literature suggests - is this why "One-Third of Juniors [are] in dark on BP tests"? (Dawe, 1993). Or is it a task that can be performed according to a set of rules as the working party of the British Hypertension Society, and the instructions in Miller and Keane, suggest?

People who are overly-familiar with a task often lose awareness of the skills and understandings required to perform that task. In order to fully appreciate these skills and understandings, it is necessary to engineer a strangeness and recapture the perspective of the novice. The next part of the fieldwork was designed to engineer this kind of strangeness, so as to clarify the nature of the task of recording the blood pressure and the skills involved. In order to take the necessary 'step back' I used the 'proxy stranger' method (see Collins, 1992b). The method involves giving a stranger basic instructions, and observing them performing the task according to these instructions. An expert native observer will be able to see the mistakes the stranger makes, and the taken for granted assumptions that they themselves hold then become clear. The advantage of this directed enquiry is that it allows those familiar with the process to recognise what they know, by observing what a stranger doesn't know. In this case it allows us to see exactly what is involved in taking the BP, and whether it can be accomplished by merely following instructions. The strangers chosen were from the sociology department at the University of Bath (4). They had no medical background or anatomy training, but all spoke English and were university educated.

3.2 The Proxy Stranger

The first volunteer was issued with a list of instructions. I created these with the intention of making them clear enough to follow and get the job done (figure 7.1). Before starting the experiment the stranger read them through and it was immediately apparent that the terminology used and her lack of anatomical knowledge were stumbling blocks. My first task was to explain the meaning of auscultation and palpation, and show her where the radial pulse and the brachial pulse were located. She did not know what 'occlude' meant in this context (5). Before she started I also suggested the best place for her to sit in relation to the 'patient'.

Her next enquiry was about which arm to use -

"Should I use the arm nearest the heart?"

I had not mentioned this in the instructions, not considering it to be relevant. Another point the volunteer raised was that she thought this test should be done after the patient had 'run around a bit'. I assured her this was not necessary for taking a resting BP. I next showed her the right way

Figure 7.1

INSTRUCTIONS FOR VOLUNTEERS IN BLOOD PRESSURE MEASURING EXPERIMENTS #1

1. Position the patient comfortably. Sitting is acceptable, with the arm resting on the arm of a chair, and the upper arm roughly level with the heart.

2. Locate the radial pulse and the brachial pulse by palpation.

3. Attach the cuff to the upper arm, above the elbow, over the brachial artery. Position it so that the two rubber tubes lie over the artery at the midline (inside the elbow).

4. Palpate the radial artery. Close the air valve above the air pump bell and pump the cuff up until the pressure that the cuff exerts is sufficient to occlude both the blood flow in the artery and the radial pulse.

5. Continue pumping to 20 or 30 mm Hg beyond this point, keeping the palpating fingers in position.

6. To find the systolic value, keep palpating at the position of the radial pulse, and slowly deflate the cuff by opening the air valve. Note the value when the radial pulse reappears. This is the systolic value.

7. At this point the Korotkoff sounds will be audible over the brachial pulse. Continue to deflate the cuff slowly whilst auscultating the brachial pulse using the stethoscope. Note the point at which the sounds disappear. This is the diastolic value.

8. Express the blood pressure as systolic over diastolic.

9. Feel free to comment on the values obtained.

to position the lever on the stethoscope so that sounds would be audible through the diaphragm. She tested it by tapping too hard with her fingernail on the diaphragm, with the stethoscope in her ears. Her reaction to this was "Ow, ow, oh shit".

The first practical task was to attach the cuff over the brachial artery. The brachial artery is in the upper arm, above the elbow, but her assumption was that it was in the same place as the brachial pulse - in the bend of the elbow. I explained that the artery is actually above the pulse point. She then attached the cuff in the right place. "[Shall I] Close the air valve?" was the next question. I showed her how to close the air valve on the pump, so that no air escapes from the cuff as it is pumped up. She tried to pump up the cuff until she could no longer feel the radial pulse. The point at which she could not feel the pulse caused some confusion. Could she or couldn't she? She wasn't sure, and deliberated about it with the cuff fully inflated. (At this point the patient complained that the veins in his arms were sticking out). She carried on and decided that the systolic value was 130 mm Hg. She next had to listen to the sounds in the brachial artery and decide when they disappeared. Confusion set in here, she tried to feel when the sounds disappeared, rather than listen. I explained that it is impossible to feel when a sound disappears, and redirected her. She was unable to hear anything through the stethoscope. When she eventually claimed to have found the right spot at which to listen, she deflated the cuff to zero, without noticing the point at which sounds disappeared and what the diastolic value was. The patient apparently had a BP of 130/0 mm Hg. By this time his fingers were tingling and white because the process had taken so long.

Volunteer number two used the same instructions, and again the location of the radial and brachial pulses and the meaning of the terms auscultation and palpation had to be explained. She proceeded, attaching the cuff back to front. Once I rectified this, she was confused about how high to pump the cuff, "'what exactly does till you occlude the blood flow mean?"

I explained that this is when you can no longer feel the pulse in the radial artery. The stranger had trouble with this -

"It is really difficult, you have the pulse, but when you start to try and concentrate on something else it is not so easy. It is difficult to know when you have lost something."

She couldn't manage the pumping up and continuing to feel the pulse in the radial artery. When she felt she had pumped sufficiently to occlude the artery, she had problems deflating and simultaneously feeling when the pulse came back in. Listening to the brachial artery sounds she found impossible - she couldn't hear the sounds through the stethoscope and gave up the procedure, fearful that she was causing pain to the volunteer patient. She couldn't do two things at once, and failed to produce either reading.

The third volunteer was issued with revised instructions which reflected the problems the first and second volunteers had encountered (figure 7.2). These new instructions explained the terms auscultation and palpation, specified where the radial and brachial pulse are found, and clarified the level to which the cuff needed to be pumped up. The instructions stressed that the brachial pulse had to be listened to through the stethoscope to find the diastolic value.

Figure 7.2

INSTRUCTIONS FOR VOLUNTEERS IN BLOOD PRESSURE MEASURING EXPERIMENTS #2

1. Position the patient comfortably. Sitting is acceptable, with the arm resting on the arm of a chair, and the upper arm roughly level with the heart.

2. For reference,locate the radial pulse inside the wrist and the brachial pulse inside the elbow by palpation (feeling).

3. Attach the cuff to the upper arm, above the elbow, over the brachial artery. Position it so that the two rubber tubes lie over the artery at the midline (inside the elbow).

4. Palpate the radial artery. Close the air valve above the air pump bell and pump the cuff up until the pressure that the cuff exerts is sufficient to occlude both the blood flow in the artery and the radial pulse. ie pump until you can no longer feel the pulse at the wrist.

5. Continue pumping to 20 or 30 mm Hg beyond this point, keeping the palpating fingers in position - where the radial pulse was evident.

6. To find the systolic value, keep palpating at the position of the radial pulse, and slowly deflate the cuff by opening the air valve. Note the value when the radial pulse reappears. This is the systolic value.

7. At this point the Korotkoff sounds will be audible - through the stethescope - over the brachial pulse. Continue to deflate the cuff slowly whilst auscultating (listening to) the brachial pulse using the stethoscope. Note the point at which the sounds disappear. This is the diastolic value.

8. Express the blood pressure as systolic over diastolic.

9. Feel free to comment on the values obtained.

This volunteer proceeded well up to the point where he had to listen to the Korotkoff sounds through the stethoscope. He was using the wrong part of the stethoscope, so I intervened to flick the switch in the right direction. Still he had no success:

"I can't hear sod all",

He gave up at this point and no diastolic value was obtained.

For the fourth volunteer the instructions were revised further so as to reflect the problems experienced by the first three volunteers (figure 7.3). I included advice on how to use the stethoscope - the flick switch should be toward the user, and the silver diaphragm should be used to listen for sounds.

Still the fourth stranger had problems - he could not find the brachial pulse until I showed him exactly where to feel, he attached the cuff inside out so that the fastenings were invisible, and he did not realise that he was the 'user' of the stethoscope and the switch should be toward him. His efforts led to him suggesting that his 'patient' had a systolic value of 95 mm Hg. He could not hear the sounds for the diastolic -

> "What is it meant to sound like? I can't hear anything."

His final comment was that the systolic value of 95 was 'pretty average', but getting the diastolic value was far more difficult because,

> "When you let the pressure out of here, it shoots down, doesn't it, with a vengeance."

His idea that 95 was a pretty average reading for the systolic was the only comment that any of the four volunteers made about the values. None of them offered an opinion about the relevance of the readings they obtained.

Figure 7.3

INSTRUCTIONS FOR VOLUNTEERS IN BLOOD PRESSURE MEASURING EXPERIMENTS #3

1. Position the patient comfortably. Sitting is acceptable, with the arm resting on the arm of a chair, and the upper arm roughly level with the heart.

2. For reference,locate the radial pulse inside the wrist and the brachial pulse inside the elbow by palpation (feeling).

3. It is necessary to become accustomed to the stethoscope. The 'flick switch' near the diaphragm must point toward the user in order for sounds to be heard through the diaphragm. With this stethoscope, the silver diaphragm is used to listen to sounds, not the black 'bell'

4. Hang the stethoscope around the neck, so that using it and inserting the ear pieces into the ears is easily achieved when needed.

5. Attach the cuff to the upper arm, above the elbow, over the brachial artery. Position it so that the two rubber tubes lie over the artery at the midline (inside the elbow).

6. Palpate the radial artery. Close the air valve above the air pump bell and pump the cuff up until the pressure that the cuff exerts is sufficient to occlude both the blood flow in the artery and the radial pulse. ie pump until you can no longer feel the pulse at the wrist.

7. Continue pumping to 20 or 30 mm Hg beyond this point, keeping the palpating fingers in position - where the radial pulse was evident.

8. To find the systolic value, keep palpating at the position of the radial pulse, and slowly deflate the cuff by opening the air valve. Note the value when the radial pulse reappears. This is the systolic value.

9. At this point the Korotkoff sounds will be audible - through the stethescope - over the brachial pulse. Continue to deflate the cuff slowly whilst auscultating (listening to) the brachial pulse using the stethoscope. Note the point at which the sounds disappear. This is the diastolic value.

10. Express the blood pressure as systolic over diastolic.

11. Feel free to comment on the values obtained.

It was clear to me that the instructions needed to be enlarged for each volunteer, and still queries and misunderstandings arose. The problems and confusion experienced by the volunteers showed me the things that experts knew about the procedure and took for granted, which the novices did not know and which I tried to make specific in a list of instructions.

They as novices did not have the necessary skills and expertise to do the job, and my instructions did not convey these adequately. Such things as knowing what auscultation and palpation mean, knowing where the brachial pulse and radial pulse are and how to find them, how hard to press, how to attach the cuff and where to attach it, how high to pump the cuff and how to do that whilst still feeling the pulse. I was able to decide when the pulse was and was not present, and I knew how to use a stethoscope, what to listen for, and when to decide that sounds were no longer present. All of these things were part of my expert 'taken for granted' knowledge of the procedure of taking the blood pressure in the traditional way. The efforts of the strangers showed how internalised or invisible these things had become to me (6). This is the expertise that has become 'hidden'. These experiments show that there is, literally, more to taking the blood pressure than meets the eye - and that there is more involved than following specific instructions and executing behaviour-specific action.

The question that arises from this is, if the task is so complex, involving many learned, internalised skills, how does the computerised machine cope so well? The machine is not part of the medical form of life - it is an asocial artefact. Answering this question involves looking at the exact task that the machine performs. What aspects of the human role have been subsumed by the machine?

3.3 The Humans' Role and The Machine's Role

3.3i The Traditional Human Role and The Skills Required

The proxy stranger experiments show the range of human skills required to take the blood pressure. Locating the pulses, attaching the cuff, inflating it, feeling, listening through the stethoscope, deciding the systolic and diastolic levels, deciding on the significance of the reading in terms of this particular patient. The human needs to master each of these procedures in order to produce a BP reading and to act appropriately in the light of that reading.

3.3ii The New Human Role and The Skills Required

When using an automatic device to monitor a patient's BP the human operator has to approach the patient, explain the procedure, attach the cuff and set the alarms on the machine. Only then does the machine do its job. And after that the human has to transfer the BP reading from the digital display or paper print out to the patient's charts and notes, and decide whether to act on the results. The parts that the human does have changed very little.

3.3iii The Machine's Role

We can see that the machine does a small part of the overall procedure - it inflates the cuff and senses the pressures as the cuff deflates. At the beginning and at the end of the process the human role remains the same as it was before the introduction of the new technology. Under the new system the machine fits nicely into the middle of the human contributions.

A doctor who has been extensively involved with automatic blood pressure measuring summed this up well:

> "All it does is save you pumping up and listening or palpating. It stops you having to go along every hour, maybe forgetting, stops the same person having to go along every hour. It does the boring bits. It is useful for things like that, but you've got to know its limitations."

The human contribution at the start of the process 'digitises' the input, narrows down the wider world of, for example, different kinds of patients, different sized patients and different sized arms, and presents it in a form that the computer can deal with. At the end, the human 'repairs' the output insofar as they decide how relevant the reading or alarm is in the particular case under scrutiny. They convert the readings into courses of action acceptable in the medical world. The doctor pointed out that the machine does the 'boring bits' in between the digitization and repair. How does it do this?? After all, people have difficulty with this part of the task. Novices often fail - as the proxy stranger experiments show - and so called 'experts' often produce readings that do not correspond with those made by other experts. What is it about this *specific* part of the process that makes it suitable for delegation to the machine?

3.4 The Specific Task for the Machine

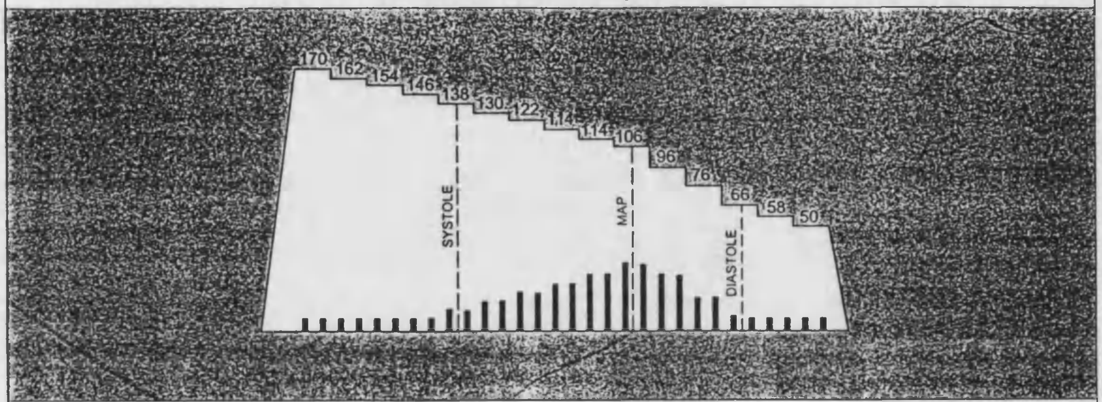
The ideal BP taking scenario would arise if all operators responded in the same way to the same sounds. If this behaviour-specific response was evident, BP readings would be comparable over time and between measurers. But, although humans may ideally want to use a behaviour-specific technique and instantiate parts of the action of recording the blood pressure with the same behaviour every time, they rarely achieve this in practice. Medical staff use various techniques, depending on who taught them and their personal preferences. Another factor is the acuity of the hearing of the operator - what one person describes as muffled, another may hear clearly. It may be inaudible to a third. When the operator listens to the pulse at the brachial artery through the stethoscope these factors come into play. Similarly, when palpating the pulse at the radial artery - who can say when a pulse is no longer present? How hard do you press? One person's BP recording is quite likely to be different to the next person's recording on the same patient. Beyond this, it is probable that readings taken by the same human vary over time, especially if they are in a hurry, or if they are tired or bored. The behaviour in each case varies - it is not as simple as the behaviour-specific model suggests. And yet it does fall into the category of behaviour-specific action because people (ideally) intend to use behaviour-specific action. (See Collins, 1990, p. 34)

The clarification of the terms behaviour-specific action and regular action, that Collins and Kusch are perfecting (see chapter 4), helps to explain how the machine can perform well and produce results that are seen as comparable over time: Collins and Kusch acknowledge that variability in instantiations occur in both regular action and behaviour-specific action. In the case of behaviour-specific acts, actors do not intend to vary their instantiations, but some degree of variation is tolerated, provided that the **outcome** of the action is 'satisfactory'. People have a margin of tolerance for variation in technique but this variation is not part of the actors' intention. In the case of BP measuring, the actors' intention - in ideal circumstances - is to reproduce earlier instantiations so that readings are comparable and observer bias is eliminated. Measuring BP is an example of behaviour-specific action where the margin of tolerance to variation in technique is relatively wide. Collins and Kusch suggest that if a machine reproduces any of the acceptable instantiations that the people use, it will be acceptable to those people. However, this case is slightly different because the machine is not reproducing any of the instantiations that a person would use, but instead uses a different method:

Forster and Turner (1986) point out that "The technique most widely used by physicians to measure blood pressure is the auscultatory method. Measurements made by this method require interpretation of the onset and disappearance of flow (Korotkoff) sounds. This procedure requires skill and is difficult to automate. In recent years there has been a renewal of interest in other methods that can be more easily automated ... the oscillometric method appears to offer the best opportunity for automation." (p. 359)

The Dinamap machine uses the oscillometric method. This was first described by Marey in 1879. The technique is an alternative to the auscultatory technique, and measures pressure changes rather than sounds (see Davis, 1982). The principle of the method is that during deflation of the cuff, the blood and the vessel wall start to oscillate when the systolic pressure is reached, and continue to oscillate until there is no constriction over the artery. The oscillations are transmitted through the air filled rubber bladder in the cuff, and measured electronically by a transducer. This procedure looks similar to the hands-on, non-invasive technique used by humans - the cuff is inflated to a level that allows measurement of the systolic pressure. Then the cuff is deflated, a further deflation step taken when two identical sequential pressure pulses are detected. This eliminates the effect of noise and ensures that accurate readings are given by the machine (Critikon publication II, p. 10). This is a sensory mechanism requiring no microphones or external transducers. The machine measures systolic, diastolic and mean arterial pressure as shown in the diagram in figure 7.4. The Critikon publication (II) suggests that "The oscillometric method of measuring blood pressure is prone to fewer errors than the conventional mercury sphygmomanometer and stethoscope technique, and observer bias is removed" (p. 11).

DETERMINATION SEQUENCE



Systolic Pressure is the cuff pressure at which oscillations in pressure begin to increase in amplitude.
The Mean Arterial Pressure is the lowest cuff pressure at which the maximum oscillations in pressure occur.
Diastolic pressure is the cuff pressure at which the oscillations stop decreasing in amplitude.

Figure 7.4

The machine is using a different technique to that used by human operators. This is where the example leaves Collins and Kusch behind. The machine is not mimicking an instantiation used by humans. But it is producing acceptable results using a different rule-based method. I suggest that if the desired outcome is produced by a machine, the method used by the machine will then fit into the humans' margin of tolerance for variations in technique. Going further than Collins and Kusch, it seems that in cases where we are indifferent to the way a task is performed, a machine does not need to mimic what any one human would do. It needs to produce an acceptable outcome. It follows that the method used will be an acceptable one. The BP machine is doing a job that people often find difficult, using a different method to people. The machine substitutes the humans' badly executed attempts at behaviour-specific action with well executed mechanical behaviour-specific action.

3.5 Advantages of A Mechanised Approach

Critikon, the manufacturers of Dinamap, have suggested that the current confusion over the 'correct' human technique has meant that "both hypotension and hypertension may be over or under diagnosed. Patients may be misclassified and treatments either withheld or prescribed inappropriately." (p. 5) They go on to claim that it is "against this background that the use of automated devices is growing rapidly." (p. 5) In practice, the machine does seem to remove the 'observer bias', as one of the nurses explained. She said that the major advantage of using the computerised technique was that it produces 'standardised' readings:

> "This machine will always give the same BP on a patient, as long as their condition is the same. Not like with people. They can vary. And different people can give different readings."

Staff had their own strategy for minimising this human variation, which a nursing sister explained:

> "When you are doing observation on, say a patient with a blood transfusion, hopefully you won't chop and change [the person

recording the BP] except at the end of a shift when those changes can be noted."

When the machine was used, there was no longer to worry about variation over time or between measurers. The machine's behaviour-specific action was more useful than the unreliable attempts of the staff to do behaviour-specific action.

4. DISCUSSION

Several points have been highlighted by the fieldwork: First, taking the blood pressure is best achieved using regular action for some parts, and behaviour-specific action for other parts. Although a behaviour-specific response would be the ideal way to do the 'detecting and measuring' parts of the job, in practice humans rarely manage this. Two types of variation occur: (1) Inter-human variation and (2) Intra-human variation.

This machine is being used to do the detecting and measuring parts of taking the BP. The machine uses a different technique to people and by-passes the subjective decisions that humans make about when the sounds become muffled, or disappear, or when pulses are no longer palpable. Consequently, the machine, which uses a standardised procedures and produces completely standardised readings gives what are considered 'better' results over a period of time than would be produced if several humans undertook the task between them. Machines are better at behaviour-specific action than we are - examples of humans choosing to instantiate action with the same behaviour every time are difficult to locate. This is despite F W Taylor's dream of routinized factories full of workers doing tasks in a pre-specified manner.

Some of the other parts of this task could also be reduced to a behaviour-specific format, and delegated to machines. Attaching the cuff to the arm is another part of the job that can in theory be mechanised in this way, without loss, because it is feasible to anticipate most variations in arm size. The desired outcome of 'attaching the cuff correctly' could be achieved using any method that 'works', which would then fall within the zone of tolerance of the humans (7). However, other parts of the process do not lend themselves to a behaviour-specific approach. They are situated near the top, wide end of the venturi. These parts of the procedure are influenced by broad, wider-world type factors and individual patient's characteristics. Explaining the procedure to the patient, setting the alarms and deciding on the relevance of a given numerical reading fall into this category. These factors cannot be reduced to a behaviour-specific format without losing a great deal of what is important in the actors' world. When people undertake this kind of task, they respond to variations in a spontaneous way. Part of the actors' intention when performing these regular acts is to give variable responses in each individual case. The outcome of a behaviour-specific approach would not be 'acceptable' for these wide-end-of-the-venturi tasks. As such, computerisation of these tasks seems highly unlikely.

The experiments highlight exactly how complex a task it is to take the BP in the traditional fashion and how many skills are involved in producing a BP reading. The machine accomplishes its part of the job - a part best done using behaviour-specific action - because we surround it with skill and expertise. This expert reaction to the prevailing circumstances amounts to digitization and repair, and it enables the machine's concerted behaviour to fit into the medical world of concerted action. The humans use their accumulated hidden expertise and commonsense abilities to fill in the gaps between the machine's level of competence and the real world of medicine. This 'hidden expertise' allows the machine to function effectively without crossing boundaries into other facets of medicine. The staff deal with these external influences when they narrow down the wider world to specifics that the machine can deal with and when they adapt the machine's output for use in the wider world. A senior registrar talked about this. He was adamant that human skill and expertise were still necessary when using an automatic machine, but this skill went largely unnoticed. He identified two specific cases. First, he explained the importance of setting the alarms at the appropriate level:

> "I've seen people use that machine to monitor blood pressure after liver biopsy, but then you have got to set it within certain limits - what level would you have it alarming at? They [the patients] could have esanguinated [bled to death] by the time it alarms if you set it low enough."

The 'correct' level at which the alarm should be set is different for every patient, because 'normal' varies for every patient. A decision based on their medical condition, their age, any investigatory procedures they may have undergone, etc, is made in each case. He then went on to talk about transferring the reading to the notes, and deciding whether to act on it. He was adamant that expertise and skill were essential at this point:

> "If the machine gives a reading of 50/20, what does a non qualified person do? Write it down, oh 50/20. Ooh, 50/20 that's alright, lovely."

Such a reading is dangerously low, but it requires expertise to recognise and act on this (8). It is necessary to be well-versed in the significance of BP readings in many different kinds of patients. This human expertise, which takes account of individual patient characteristics, is essential if the machine is to fit into the social setting.

5. SUMMARY

The theory that machines can reproduce that part of human action which people choose to carry out in a behaviour-specific manner is not challenged by this study. An interesting additional point illustrated here is that machines using one set of behaviour-specific responses can be used in place of humans who have failed in their attempts at a different set of behaviour-specific responses. The BP machine's results are as acceptable as the human results, (or, in this case, more acceptable) even though the method used to do the task is different.

This machine does its job in a way that is completely acceptable to the users - it is well received and blends into the social setting of wards. The users contribute skills that enable this blending-in to occur, but the tedious, methodical part has been handed over to an automated device. There seem to be no barriers to the use of this machine, either human barriers or technological barriers. This is because this is a very limited field of medicine, not affected by external influences, where behaviour-specific responses are the best way to get the job done. This is a task that sits in the narrowest part of the venturi. Employed at this level, the machine works well. Do any other facets of medicine fall into this category? Does the success of the automatic BP measuring machine occur in other instances? The next case study involved investigations into the use of automatic machines designed for interpreting electrocardiograph traces. Does this case contribute anything new to the theory of behaviour-specific action, or tell us anything new about the usefulness of the venturi model?

FOOTNOTES TO CHAPTER 7

1. Sometimes the point at which the sound of the blood flow muffles is used to determine diastolic pressure. There is much debate about which point should be used. Some textbooks suggest that diastolic pressure should be recorded at the 4th phase (when the sounds muffle) eg. see Macloed (ed) 1979, p. 112. Other texts suggest phase 5 (when the sounds disappear) eg. see Perloff 1982, p. 50; Miller and Keane 1978, p. 132.

2. This is an important claim because it is generally accepted within cardiology that an intra-arterial pressure value is the next best thing to a direct pressure from inside the heart. To be more accurate than an intra-arterial pressure, the value would have to be recorded within the chamber of the left ventricle using a cardiac catheter.

3. Gaining native competence and understanding during this case study was easy. I drew on my own experience as a cardiology technician familiar with both methods of BP measuring and with the significance of the readings, and I drew on the experience of the experts involved with the new technology. When full participation was not possible, I used conversations with users and periods of observation of the machine in use to see how the machine performed and how well it fitted into the routine of the department.

4. A problem with this method is deciding how strange the stranger should be. An extreme choice of stranger would be one so strange that she did not speak the same language as the natives. This would illustrate all the expertise to do with understanding English that is involved in the task at hand. The stranger need be as strange as the experimenter chooses. In this case my choice was based on the need for someone unfamiliar with the medical task under scrutiny, but familiar with other aspects such as English language knowledge, which are needed for the job.

5. In retrospect it may seem that I was naive to have expected novices to the medical world to understand these anatomical terms, and that including them in the instructions was disingenious. At the time of formulating the instructions these terms seemed perfectly appropriate. This may signify how invisible the extent of my medical socialisation had become to me. As such it reinforces my claim that a step back needs to be taken by those overly-familiar with any task.

6. They were so familiar to me that I needed someone else to show them to me. A poet once sumed this up:

> "You need someone to open up a door,
> To show you something you seen before,
> but overlooked a hundred times or more.
> You need someone to open your eyes."

(Bob Dylan, 1963)

7. A senior registrar had suggested that attaching the cuff to the arm was one of the more skilled parts of the task. He asked me, "Well would you know how to fix it to the arm? Would anybody else know? Would my mother know how to fix it to the arm over the brachial artery?" He was stressing the expertise needed for this part of the procedure, and would presumably argue that this part could not be delegated to a machine. My

suggestion is that anticipating the likely variations associated with this part of the procedure is more feasible than anticipating the likely variations associated with other parts of the task, so an acceptable mechanically produced outcome is more easily envisaged. The necessity for an experienced human to stand by and intervene when a patient with an arm type (or without a left arm) that has not been anticipated is still necessary.

8. This is a very broad claim. Is this kind of knowledge really limited to experts? An ad-hoc poll amongst my non-medical peers suggests that knowledge about blood pressure readings is indeed esoteric. When asked what a reading of 50/20 meant to them, the responses varied. Many had no idea, being unaware of what a normal reading was. One person said, "That sounds OK to me, because isn't the bottom figure supposed to divide into the top figure 2 or 3 times?"

SECTION 2

CHAPTER 8

INTERPRETIVE ELECTROCARDIOGRAPH MACHINES

CHAPTER 8 - INTERPRETIVE ELECTROCARDIOGRAPH MACHINES

1. BACKGROUND, PROCEDURES AND MEDICAL TERMS

Heart attacks (infarcts), high blood pressure (hypertension), and episodes of chest pain on exertion (angina) are common-place in the developed western world. Cardiology is the branch of medicine that deals with these problems and one of the most simple Cardiac tests is an Electrocardiograph (ECG). An ECG is a graphical record of the electrical potentials produced in association with the contraction and relaxation of the Cardiac muscle. Electrodes attached to the surface of the body detect this electrical activity which is presented either on a VDU or as a permanent record on paper. Irregularities in the timing (rhythm) of the heart's complexes as well as in the shape (morphology) of the complexes are evident on an ECG trace. These cues enable an experienced practitioner to ascertain the condition of the Cardiac muscle and the electrical conduction mechanism which triggers the heart's contractions, and to tell whether a previous infarction has occurred.

Each time the heart beats, one 'cycle' is produced on the graph. A schematic cycle is shown in figure 8.1. The p wave corresponds to atrial contraction, the qrs complex to ventricular contraction and the T wave to ventricular relaxation. In normal Sinus Rhythm each p wave is followed by one qrs complex and one T wave. Deviations from this pattern signify an unusual rhythm.

Ten electrodes are positioned on the body during the recording of an adult ECG - one on each arm and leg, and six more on the chest. Various combinations of these ten electrodes produce a 'twelve lead' ECG. This is a chart showing twelve sections of Cardiac rhythm, which offers an all-round view of the heart. Most recordings have a long 'rhythm strip' printed beneath the basic twelve leads which allows detailed examination of the predominant Cardiac rhythm.

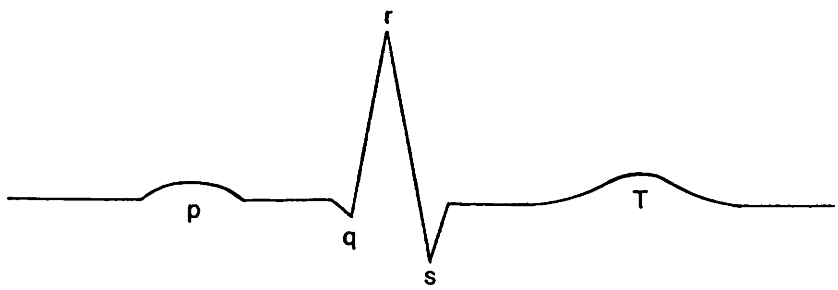


Figure 8.1 Normal sinus rhythm

p wave = Atrial depolarisation

qrs = Ventricular depolarisation

T wave = Ventricular Repolarisation

Routine ECGs are recorded on all types of patients, even those without obvious Cardiac problems. Most patients requiring a general anaesthetic can expect to have an ECG taken in advance to check for Cardiology problems, and to determine whether they are fit to withstand anaesthetic and surgery. These 'pre-op' ECGs become more likely with increased age, when the patient is a heavy smoker, or falls into some other high risk category. If a long stay in hospital is anticipated, a series of ECGs may be recorded, so that changes in Cardiac condition can be monitored. ECGs are quick and easy to perform, painless and inexpensive, and it has been estimated that over two hundred million ECGs are recorded annually world-wide (Banta et al 1985, p. 23).

Writers such as Doue and Vallance believe that "ECG interpretation (is) a complex and hard learned profession" (1985, p. 29). However, some conditions are more easy to identify than others, and certain ECGs can be interpreted by utilising knowledge about the expected normal physiology of the heart and the conduction mechanism. The 'rhythm strip' is used extensively at this level of analysis. The heart rate, rhythm disturbances and first, second and third degree heart blocks are easily seen on long rhythm strips (1).

Other abnormal ECGs are less easy to interpret and an analysis of all 12 leads of the trace is required. The patterns associated with new, old and established infarctions, ischemia or muscle hypertrophy, supra-ventricular tachycardias and axis deviation may occur alongside the rhythm disturbances described above. A basic knowledge of the physiology and conduction mechanism of the heart is necessary but not sufficient for analysing these traces - complex criteria have to be utilised and delicate patterns and trends distinguished.

This can tax medical staff for a number of reasons. When presented with these complex traces many medical staff are unable to offer an interpretation; even if they once understood the traces, nurses, medical students and doctors working in specialities other than Cardiology can easily become 'rusty'. Some staff may never have even mastered ECG interpretation because of a lack of available expert tuition, and in any case the standards of ECG analysis vary widely between institutions. To add to the confusion, it has been admitted that the parameters used to interpret ECG information are not standardised throughout the medical world. As Ginzton and Laks acknowledge, "Cardiologists do not apply criteria consistently when reading ECGs." (1984, p. 40).

2. THE COMPUTERISED MACHINE

Manufacturers of Cardiac monitoring equipment have recognised the esoteric nature of ECG interpretation skill and they see a machine which can accomplish ECG interpretation and disseminate the information to a wider audience as commercially desirable. The initial ECG analysis programs, developed in the 1960s, were greeted with excitement. Each new generation of improved machines has been hailed as a step in the right direction. Ginzton and Laks (1984) suggested that computerised ECG analysis was advantageous because of its speed, consistency and the reduction in turn-around time (the time between taking the ECG, getting a report and initiating the appropriate action) that it engenders. Other authors believe that "these analysis programs are attaining widespread acceptance as a means to help contain the cost of health care," and that they make the physicians job easier, eliminate much of the drudgery involved in ECG analysis, and allow the physician more time to concentrate on patient care. (Banta et al 1985, p. 23/24).

The justification for introducing interpretive ECG machines is that they can reduce the workload of the physician by performing some of the more tedious aspects of routine ECG analysis. It has been suggested that when people interpret ECGs, "The human expert's task can be divided into two parts; Screening - given an ECG readout, determine whether the patient is suffering any heart malfunction at all. Diagnosis - given an abnormal ECG, give a hypothesis diagnosis of the heart disorder it represents." (Priest 1989, p. 39/40). One of the major claims is that automatic machines can take over the 'screening' of ECGs. A recent study on the efficacy of interpretive ECG machines concluded that "the machines tested appear to have a place in sorting electrocardiograms into normal and abnormal [screening]." (Lack et al 1989, p. 24). This is a point that is emphasised by manufacturers of the automatic machines - if the ECGs are screened automatically, the physicians have the luxury of only over-reading those traces deemed 'abnormal' by the machine. This gives them more time to do other, more interesting things. A sales representative explained it as follows:

> "We are trying to limit the amount of work the Cardiologist is having to do ... They are not interested in reading ECGs any more if the machine can do it and they can just look at a few."

Beyond this, the machines are designed to measure various sections of the ECG and to offer an interpretation of all ECGs, including those deemed 'abnormal'.

3. THE FIELDWORK

The fieldwork was carried out at three hospitals in Wales. Between 1981 and 1986 I had worked permanently in the cardiology department at one of these hospitals and participant comprehension was thus easily achieved. My prior experience and the insider understandings I held meant that I was completely at home in the world of cardiology. I knew what was and what was not admissible as an account within the field and I experienced no language barriers - in terms of accent or technical jargon. Dealing with staff of all grades posed no problems in terms of misplaced deference. I was fully aware of the accepted and expected conduct within the hospital, I was interested in new developments in the field and was able to engage in authentic two-way exchanges with the professionals and patients involved. Most important for this study, I learned my ECG interpretation trade at a centre of excellence, and am able to analyse most ECGs. I visited one department every day for several weeks, assisting, talking and collecting examples. At the smaller department I spoke with all the staff in the department, from technicians to once-a-week consultant, and observed as the machine was used. At the third hospital, I was given a uniform and allowed to do the job of a technician for a month. The objectives were to assess the nature of the task of ECG interpretation, is it regular action or behaviour-specific action, and where in the ventrui does the task fit? How well do interpretive ECG machines perform, and how well do they fit into the social/medical environment? Do the answers to these questions alter in view of the particular medical environment in which the machine is deployed?

3.1 Day To Day Use

My first impression was that the staff foresaw no problems in allowing the machine to screen the ECGs. Theoretically they saw it as a straightforward job that a machine should be able to do. For example, one registrar specialising in Cardiology suggested that,

> "An ECG is just a line on a piece of paper, so all you are doing is describing what you see, based on a background of what you believe to be normal, and only then deciding whether it's abnormal or not ... given the technology, a machine should be able to say whether a black line on a white piece of paper is normal or not."

As long as the machine limited itself to describing the ECG in relation to what it is programmed to define as normal he foresaw few problems. A second expert stated that

> "I think that you can clearly get a machine that will tell you if an ECG is normal - I don't think that's any problem at all."

Many of the staff were willing to rely on machines to carry out the screening function, and a concensus emerged along the lines of:

> "In cases where the machine says it is normal, then fine, I think we can rely on that."

But when I looked at the machine in practice it became clear that in everyday use practitioners have become accustomed to machine failures and mis-interpretations. One member of staff recounted an instance that had occurred a few days previously, which involved a patient with third degree heart block. Such patients are usually assessed quickly with a view to pacemaker implantation. The automatic machine had analysed the ECG and made an interpretation of 'normal sinus rhythm'. This mistake was serious, and the technician summed up her reaction to the mistake:

> "Its hard to believe that it could get something so basic, wrong. I think it's because in this instance, the rate was fast for third degree block - it was about 60 - so instead of looking for the p waves it just assumed that it was normal sinus rhythm because it was so fast. I think it has real trouble sensing the p waves, and was only sensing the qrs."

> "So what did you do about it?"

"I did another ECG with the interpretation switched off. You can't have them hanging around with false diagnoses on them."

This example of a 'false normal' was the most severe that the technician could remember. Soon though, an instance where the machine produced a 'false abnormal' occurred: In the ECG reproduced below (figure 8.2), the machine has produced an interpretation that includes atrial flutter/fibrillation. It has 'detected' multiple p waves at a rate of 343 bpm, a ventricular rate of 62 bpm, and declared the ECG to be abnormal.

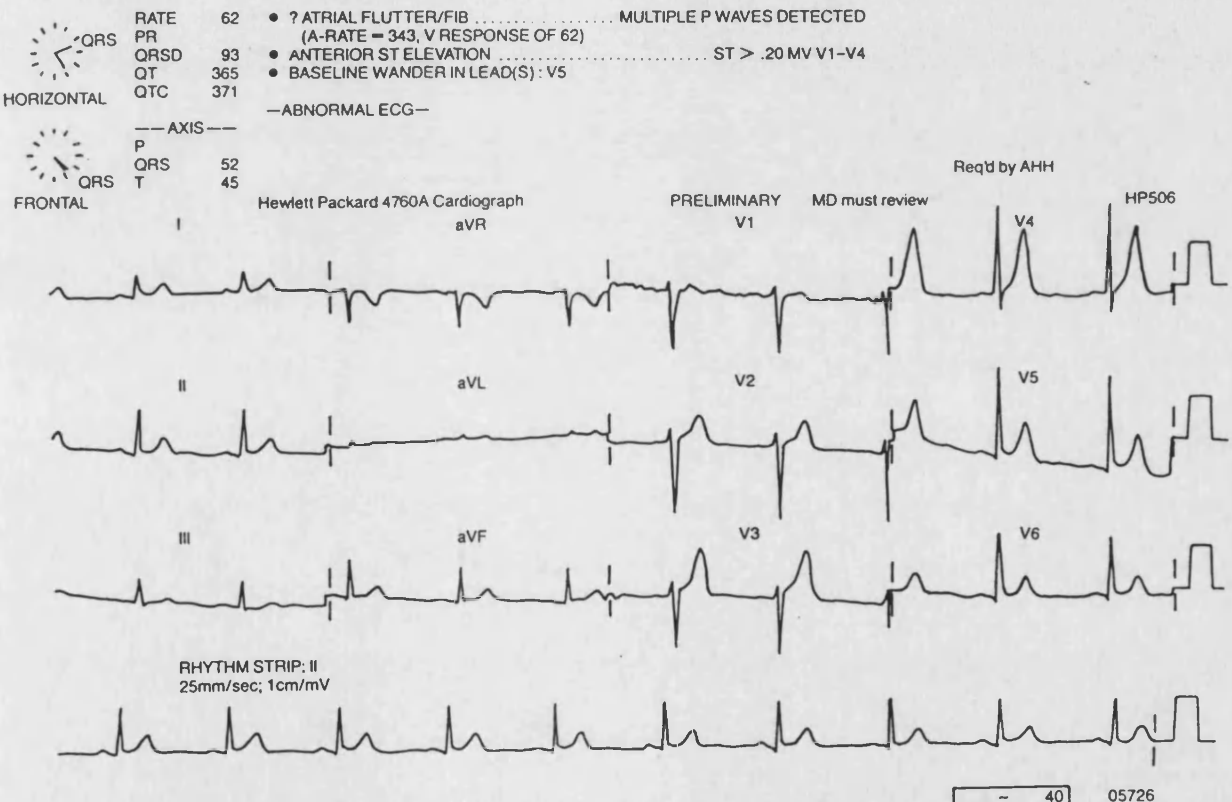


Figure 8.2 Example of false abnormal reading

This interpretation is presumably based on the pattern evident in lead V1. But it is <u>not</u> possible for an ECG to show atrial flutter or fibrillation in just one of the 12 leads of the ECG. In this situation a competent human analyst would view the ECG as a whole. This would provide a resource for finding the ECG to be of good quality except for lead V1, and they would subsequently be able to make an analysis based on the information in the other eleven leads. The result would be an interpretation along the lines of:

- Loose electrode or interference in V1,
- Normal sinus rhythm at 62 bpm,

In both these cases the machine has failed to separate the normal ECG examples from the abnormal (screen the ECGs) in an acceptable fashion.

What about interpreting abnormal ECGs? Is the machine better suited to this part of the job? Staff generally felt that this task was beyond the scope of computerised ECG machines, as machines do not have access to the strategies that humans use. When experts were asked about these strategies they often found it difficult to express them. The cardiology registrar said:

> "I look at the rhythm, I look at the rate, I look at the axis, I do this, I do that, but sometimes I just pick it up and say 'my God, complete heart block' and put it down again, or say 'left bundle branch block', and you don't say much more about it."

And a Cardiac Consultant stated that

> "ECG analysis is such an incredibly complex thing. It's not just what you learn in the books, it's the twenty years of experience ... knowing which T waves which are inverted are the ones that are likely to be troublesome and the ones that are not. So that experience side of things is very difficult to verbalise and very difficult to program."

Thus, it seems that human experts rely on past experience to interpret ECGs. However, whilst they know how to do it, they cannot always explain how they do it. That is, experts find it difficult to specify the characteristics of this essential tacit knowledge. Clinical judgement allows experts to decide how to deal with subtle variations in each new case, but programming this information into a machine is by no means straightforward. One senior registrar felt very strongly that a machine should not attempt to interpret abnormal ECGs, because

> "a machine can't talk to the patient and get a history, and the machine can't examine the patient."

Is this pessimistic view justified in terms of the machine's practical results? What happens when the machine interprets the ECG traces? During the fieldwork I collected numerous examples where the machine produced unacceptable interpretations. Example 3 overleaf shows an ECG with a ventricular ectopic. This was the first ECG I saw the machine interpret (figure 8.3).

The interpretation is that the aberrant beat is an atrial ectopic rather than a ventricular ectopic. This is not a life-threatening mistake - neither variety of ectopic is dangerous if it occurs singularly and infrequently. However, this mistake is interesting because it is so basic. The characteristics of an atrial ectopic are quite different to those of a ventricular ectopic. The machine has failed to interpret the ECG correctly and has failed to recognise the distinct characteristics of the most common abnormal beat - the ventricular ectopic.

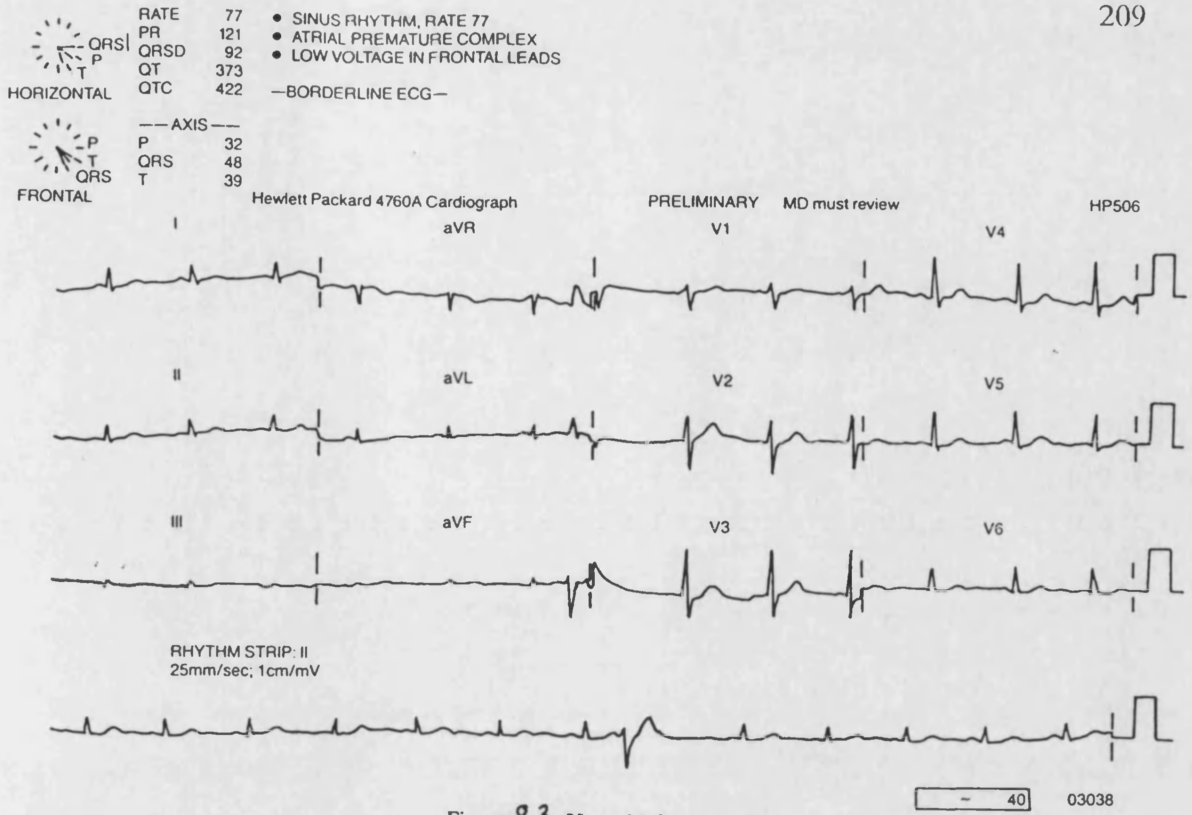


Figure 8·3 Ventricular ectopic

Another classic example of the machine's incompetence soon occurred. An ECG was recorded on a sixty nine year old man. The machine produced an interpretation of sinus rhythm with first degree AV block - which it categorised as abnormal - at 15.01 on May 25 (example 4, figure 8.4). This interpretation is inaccurate, as is the suggestion that the patient presented right-axis deviation. After some discussion with the technician in charge a second ECG was recorded, with no adjustments to the recording leads. Example 5 - the trace in figure 8.5 - was produced at 15.06, and this time the machine suggested that the ECG was borderline, showing a non-specific conduction delay. No mention now of right-axis deviation. This interpretation is also inaccurate. Finally, example 6, shown in figure 8.6, was produced at 15.13. This time the machine gave a more accurate interpretation - atrial flutter with a 4:1 block. For the first time, the possibility of right bundle branch block was suggested in the analysis.

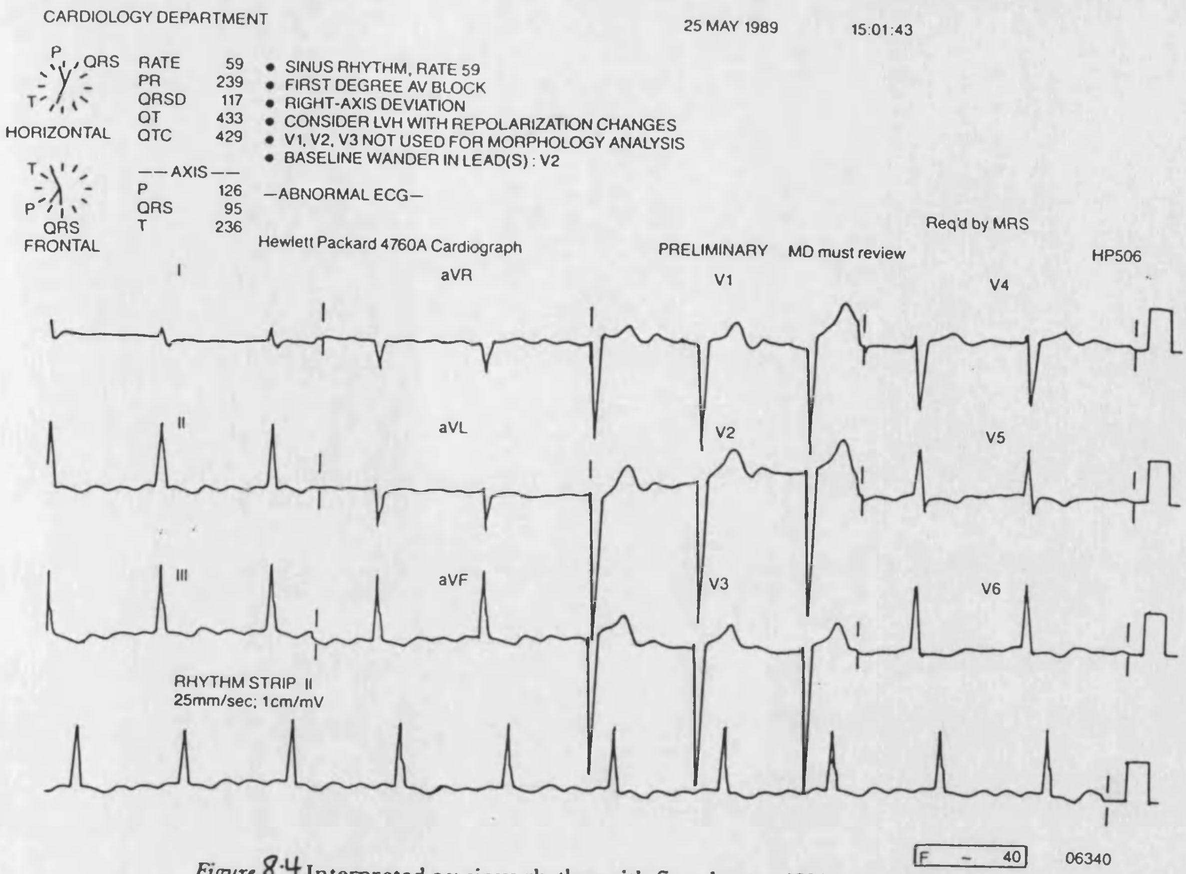


Figure 8·4 Interpreted as: sinus rhythm with first degree AV block – 15:01 hrs

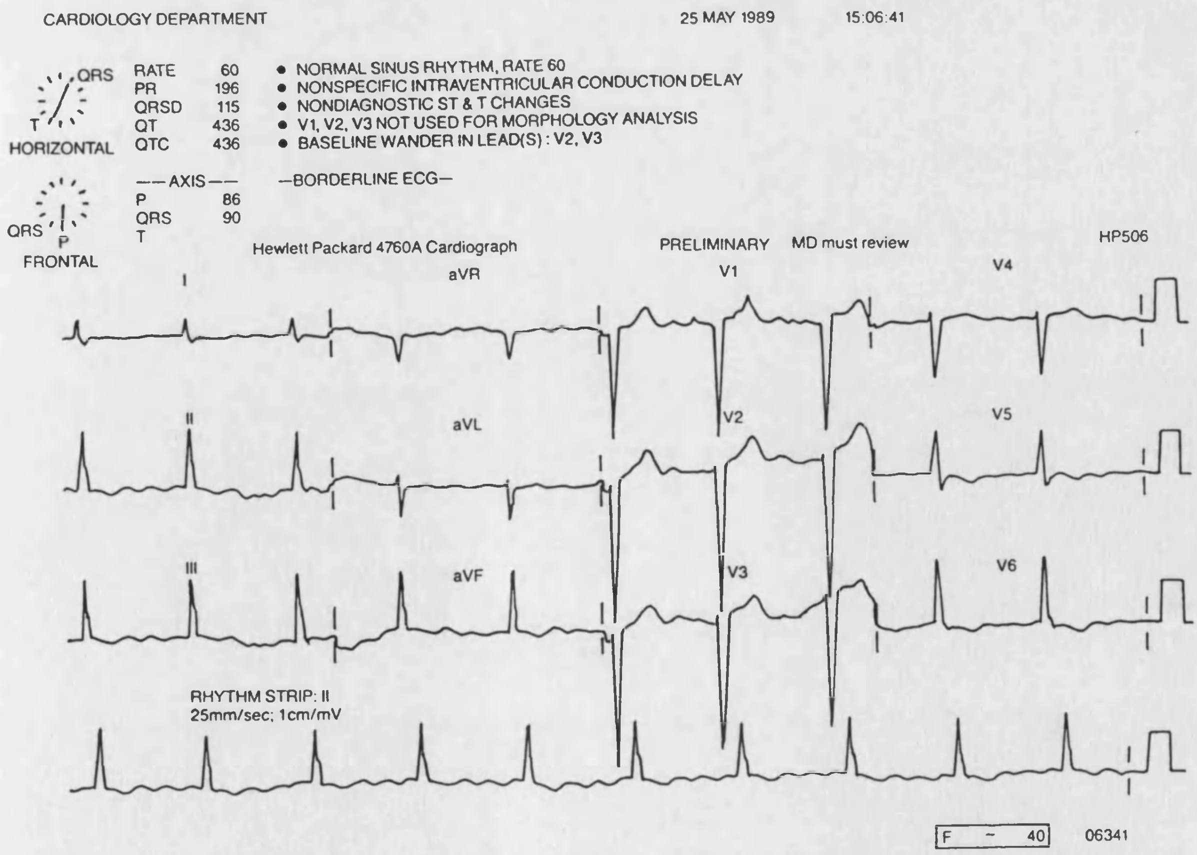


Figure 8·5 Interpreted as: sinus rhythm with non-specific intraventricular conduction delay – 15:06 hrs

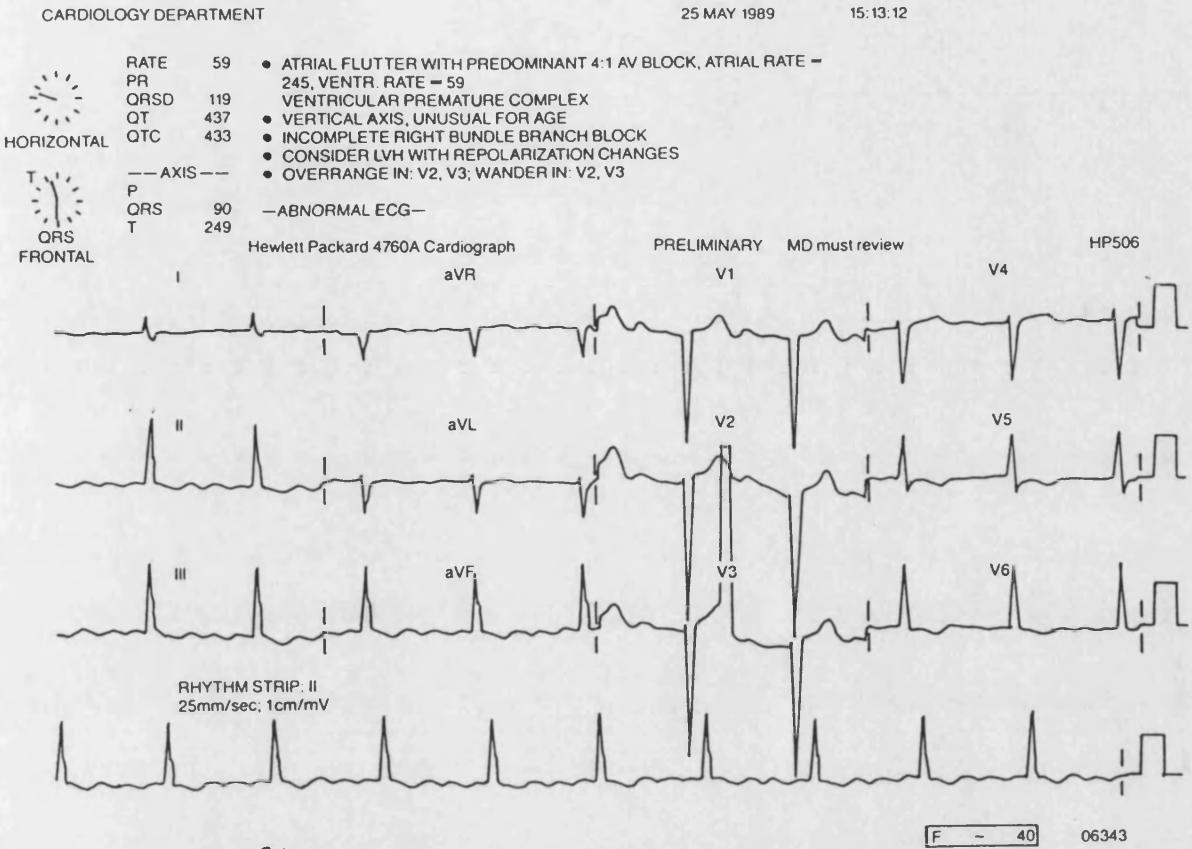
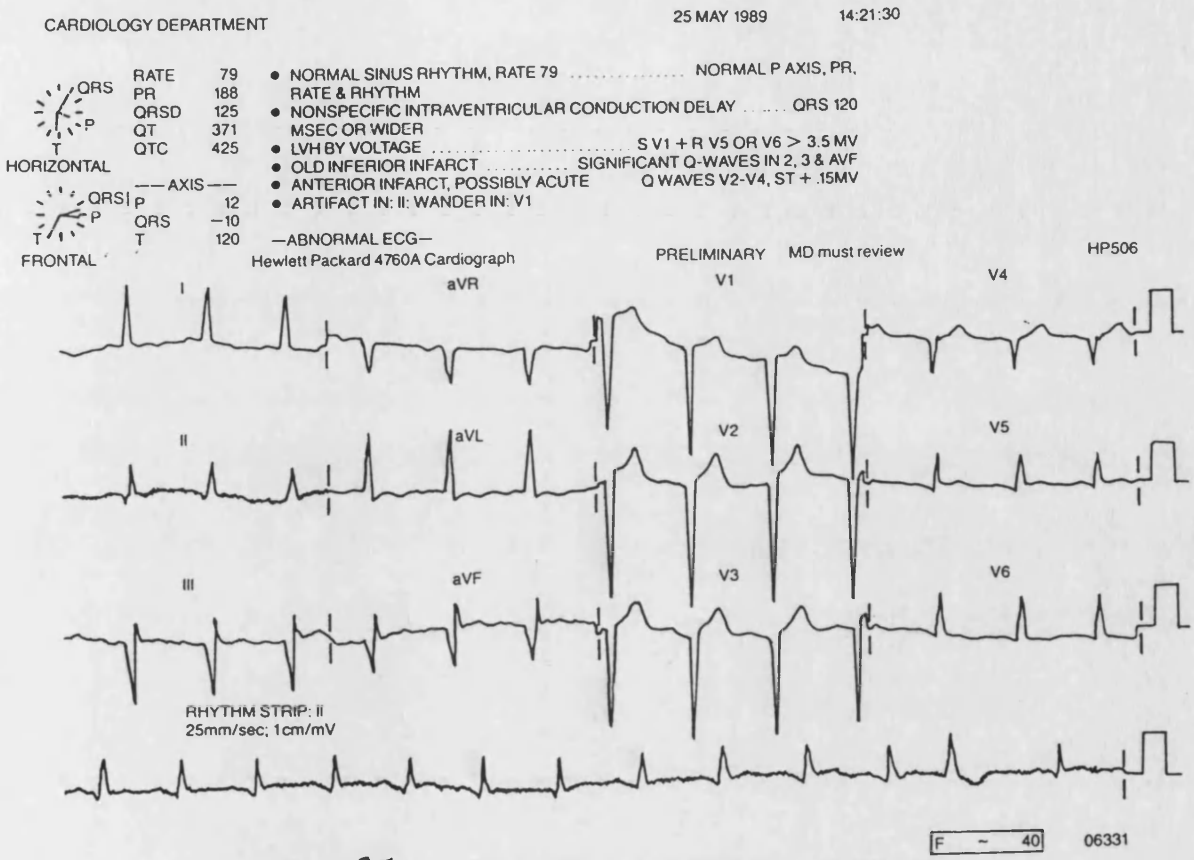


Figure 8·b Interpreted as: atrial flutter with predominant 4:1 AV block – 15:13 hrs

In this instance the machine has offered three different interpretations of ECGs taken from the same patient over a period of just twelve minutes. (These three cases are discussed in more detail in section 3.3, below).

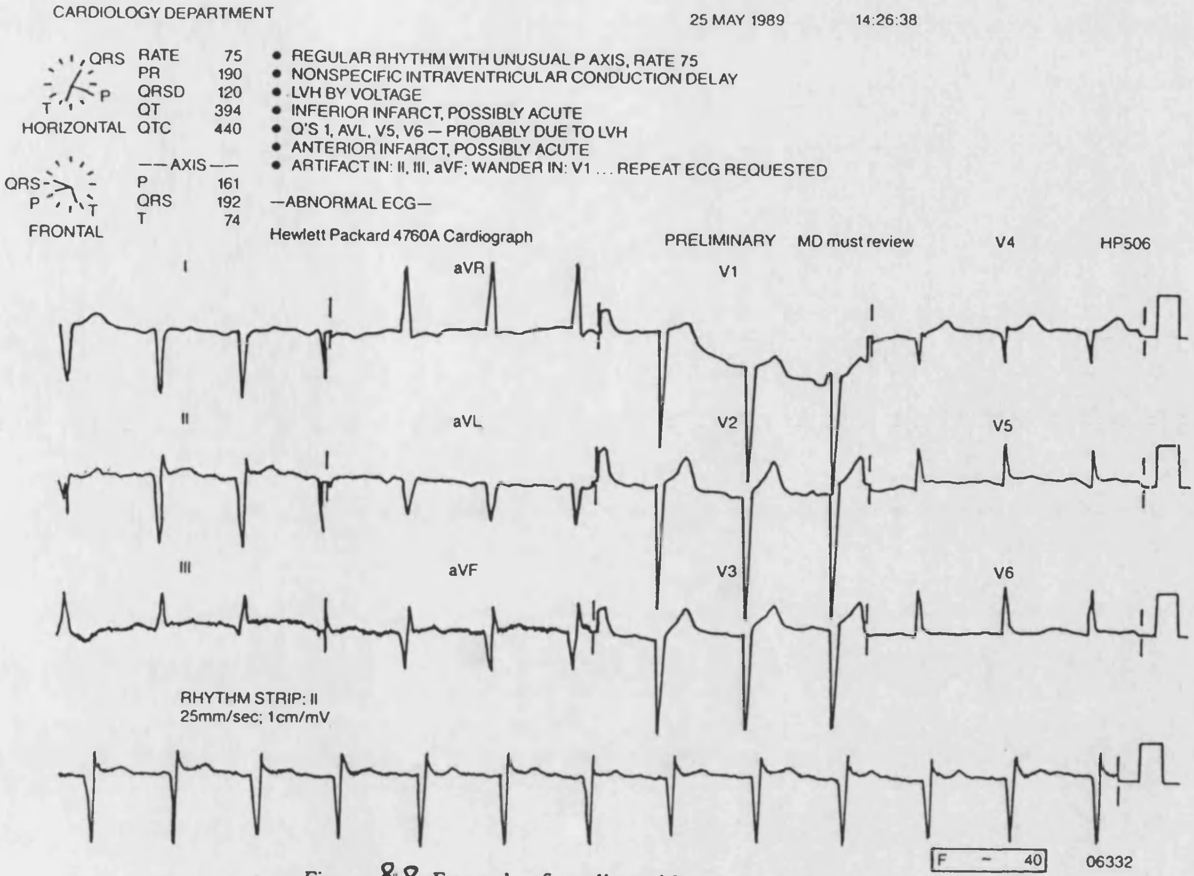
Occasionally technicians make mistakes when applying the electrodes to the patient. If the arm leads are reversed, or a leg and an arm lead are swapped, the ECG will display unusual characteristics. These are normally noticed by the technician when the ECG is printed and the electrodes rearranged so that another ECG can be recorded. However, the interpretive machine does not always detect human error in limb lead application, as is shown in the following experimental case.

CASE 1: Two ECGs were recorded - the first with normal limb lead application, (figure 8.7) and the second with the arm leads reversed (figure 8.8).



Figures 8·7 Example of reading of normal limb lead application

The trace in figure 8.7, which was recorded correctly, has been interpreted as abnormal by the machine. Figure 8.8 also has a list of abnormalities printed on it, and an interpretation of abnormal. The machine has not noticed that the trace in figure 8.8 has been recorded with the arm leads reversed (2). More of these experiments were performed on several patients. The machine did occasionally suggest that the limb lead application was incorrect, but far more often it attempted an interpretation on a trace that was recorded from wrongly positioned electrodes.



Figures 8.8 Example of reading with arm leads reversed

The details of these examples are not of prime importance here. The major point that I am making is that the machine is unreliable in its ability to distinguish acceptable input from unacceptable. It produces an interpretation on unacceptable data - and relies on the human operator to provide it with acceptable data.

In the ECG reproduced in figure 8.9 the machine has made an abnormal interpretation. A pacemaker rhythm has also been identified, but no pacemaker activity is actually present. To an experienced eye, the mistake can be explained because the qrs complexes are regular and of high voltage - similar to pacemaker spikes. The operator in this case ignored both the misinterpretation and the extraneous suggestion that a pacemaker rhythm was evident.

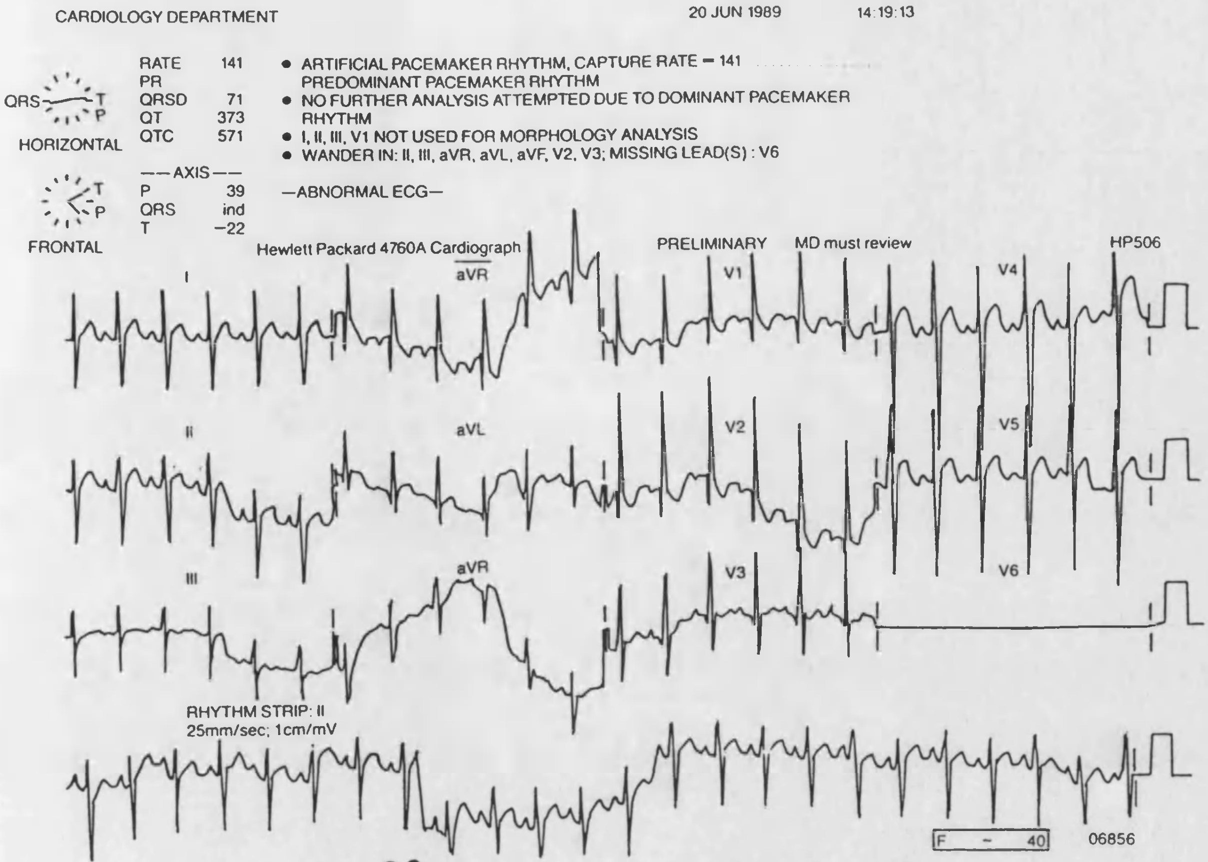


Figure 8.9 Example of machine misreading rhythm strip (a)

The ECG shown in example 10 (figure 8.10) has been interpreted as abnormal, but the essential abnormality - that of a malfunctioning pacemaker - has not been listed. This omission is serious, the patient is liable to collapse at any moment. The lack of ventricular activity on the rhythm strip shows that spontaneous ventricular contraction cannot be relied on. The pacemaker spikes on this trace are not followed by ventricular contractions, and the machine has mistakenly identified these redundant electrical pacemaker impulses as high voltage qrs complexes - suggestive of left ventricular hypertrophy. All in all, this interpretation misses the vital characteristics that humans would notice and use to categorise this ECG as dangerously abnormal. The operator reacted by recognising the mistake, mentally inserting the missing portion of the interpretation, and ensuring that it was brought to the attention of a physician.

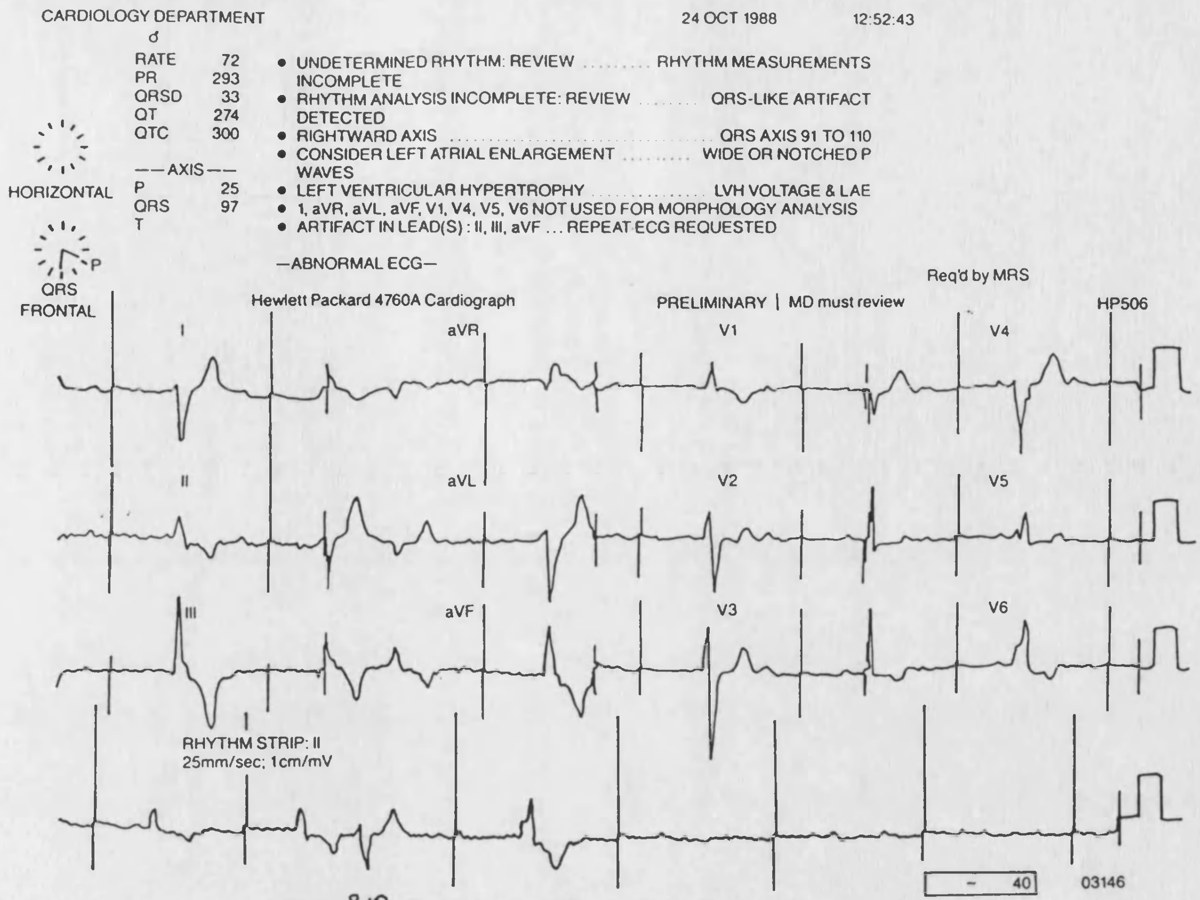


Figure 8·10 Example of machine misreading rhythm strip (b)

The examples of the machines' mistakes described here took many working hours to collect. These examples were carefully selected from many cases where the machine did give what was counted as accurate screening and interpretation decisions. In practice the machine coped well with many straightforward traces. Beyond this, the amplitude and voltage measurements produced were believed to be extremely useful in clinical settings, during drug trials, and most importantly for measurements on heart transplant patients' traces. The machine is very useful in these areas. However, the cases described above show that fully acceptable automatic ECG interpretation has not been achieved. Programs often fail to produce interpretations that equate with human interpretations. Despite this, the machines are appearing in more and more departments and apparently being used regularly. How can we explain this paradox?

3.2 Digitization and repair

First of all, staff are aware that the machines require 'clean' input data, devoid of mistakes or fuzzy data. The staff provide this 'digitized' input. When the machine is left to cope with input signals that have not been adequately digitized, such as in example 2, the output they provide is not acceptable. Examples 7 and 8 show what happens when the machine is fed faulty input - it produces interpretations on ECGs which humans would recognise as having been recorded incorrectly. If the staff do not provide an acceptable and clear input signal, the output is often unacceptable. One of the consultants was well aware of this, and told me:

> "You know the American expression 'Garbage in - garbage out'? Well, that's what happens with this machine. If you screw it up by putting the wrong stuff in, you get the wrong stuff out. It is not that intelligent."

Staff also respond when the machine produces unacceptable output interpretations: When presented with the ECG described above, which the machine described as showing sinus rhythm when third degree block was evident, the technician ignored the machine's 'normal' decision, over-read the ECG herself and decided that the machine had made a serious error. She then produced another recording without an interpretation. Again in ECG example 2, the technician immediately 'repaired' the machine's faulty output. She realised that the interpretation was wrong, adjusted the V1 electrode so as to receive an improved signal and recorded another trace - this time with the interpretation mode switched off. Similarly in examples 9 and 10.

Without human repair of erroneous output, the consequences in either of these cases may have been serious.

The concepts of digitization and repair do help explain how the machine is able to fit into the environment despite its obvious shortcomings. But how can the human reaction to the series of mistakes associated with the traces in figures 8.4, 8.5 and 8.6 be explained?

3.3 The 'Same' Traces

From a cardiological point of view the ECGs in figures 8.4, 8.5 and 8.6 are the same, in terms of rhythm, rate and morphology. When technicians and physicians were asked to comment on these ECGs the responses were as follows -

> "Ha! Identical. The rhythm is the same, the qrs morphology is the same. It is identical. These three are identical ECGs."
>
> "They are the same ECG basically."
>
> "Yes, same, yes."
>
> "I would say that they are probably all the same. I haven't measured the things, but they look all the same to me."
>
> "Yes, I can't really see any difference."

All the experts that were approached considered the ECGs to be the same. Yet finding them 'the same' is an achievement on the part of the practitioners, since the traces are in fact slightly different: the ventricular rate varies, the voltage of the complexes varies, and there is artifact on figure 8.6 in lead V3. Yet to experienced practitioners these ECGs are sufficiently 'the same' in terms of the important criteria to be classified as the same - they ignore the differences (3). Thus, concepts of similarity and difference do not come 'ready packaged' so to speak, but are organised in the context of their situated application. What people see as 'the same' or 'different' is a product of their expertise and experience. Cardiologists are able to respond to these three ECGs in the same way, in the knowledge that other Cardiologists will also respond to them in that same way, because of their Cardiological socialisation and training. The machine, though, is programmed to recognise specific differences. So, differences that the experienced human ignores cause the machine to come to, what is in practice, a wrong conclusion. The human technician displayed great tolerance and charitability toward the machine. She re-pressed the start button every few minutes until a reasonable interpretation of the ECG was produced by the machine. Had I not been there, expressing an interest in the mistakes, she said that she would have switched the interpretation facility off after the first mis-interpretation, in order to avoid dealing with the mistake. She would have spared the machine the close inspection I insisted on.

4. DISCUSSION

The commonly expressed belief was that the machines should be able to screen ECGs, but would have difficulty interpreting abnormal ECGs. The fieldwork has shown that in practice the machine faces difficulties with both tasks.

The belief that the machine can distinguish between normal and abnormal ECGs is based on the notion of a definite dividing line between these categories. But this ignores the possibility that in practice 'normal' is an achieved rather than a given characteristic of an ECG. Categorising an ECG as 'normal' involves negotiations performed the light of prevailing circumstances. This can be seen in three ways:

First, when doctors or technicians classify an ECG as normal a number of different contextual factors are taken into account, such as the age, race, medical history and body size of the patient. Thus, what is 'normal' for one population would not be 'normal' for another. What is 'normal' for an eighteen year old would not be 'normal' for an eighty year old. Attributions of 'normality' are thus based upon the medical practitioners experience of the relevant categories that can be assigned to a patient and the contextual relevance of their information. As one doctor pointed out,

> "There are some people who's body shape is so extreme that what is normal in them would be thought of as abnormal if you didn't know what the shape of their bodies were."

Second, rather than being a statistical term, related to the exact morphology of the complexes, 'normal' may refer to a state in which the patient is fit and healthy. An asymptomatic patient may present an ECG that is statistically abnormal, but which experienced practitioners count as 'normal' in that particular case. One of the consultants explained that

> "Different races and different ethnic groups have different normals."

Third, disagreement about the criteria for an abnormal ECG is widespread. Similarly, what constitutes a normal ECG is a source of debate amongst medical practitioners. As one Cardiologist put it:

> "There are as many definitions of what's normal as there are Cardiologists."

A senior registrar reinforced this point when he suggested that the idea of 'normal' changes as more information becomes available to the analyst, and they become more confident. As this confidence grows, they become able to say,

> "This, although it is a bit odd is normal in the absence of any other problem, whereas a machine does not have the confidence because its not built in to have confidence to say - 'Well, on balance this is probably just nothing, and is just a q wave in V2, maybe the heart is just a bit rotated - forget it.'"

A cardiology consultant also expressed concern that problems may arise if the machine was asked to interpret subtle variations on normal, and determine what these meant clinically in each individual case. He was wary of the machine's tendency to produce

> "worrying reports ... on what are actually normal variants."

Another doctor reinforced this when he said,

> "I think sometimes, an ECG that I would consider to be normal, the machine would come up with a list of possible abnormalities."

These points suggest that a definition of normal can only be 'for all practical purposes', and that any blanket formalisation of normal has to be moulded to fit the circumstances in which it is used. A consequence of this is that the meaning of normal and the position of the normal boundaries can be negotiated, and when human analysts disagree, considered discussion leads to a consensus interpretation. 'Correct' interpretations of ECGs, and the normal or abnormal decision are achieved by the participants - the physicians and technicians involved - using regular (mental) action. Pre-defining normal is difficult, especially in borderline cases. Consequently, programming machines to identify normal cases is a dubious concept (4).

Similarly, when interpreting abnormal ECGs the machine can make quite drastic mistakes. This is because just as the classification of an ECG as 'normal' or 'the same as' or different from another trace is a situated achievement, so too is the classification of a trace as 'abnormal'. During ECG interpretation the machine applies criteria in a pre-specified manner which does not match the way that human experts interpret ECGs. Machine interpretations which differ from humans' interpretations arise because the machines do not approach the task in the same way as humans do, or use the same cues.

Another problem with the machine is that it often 'over-diagnoses', and interprets what people would consider to be normal ECGs as 'abnormal'. One consultant suggested that this occurs because:

> "The philosophy of the machine is to diagnose normality with certainty. The corollary of that is that it must also diagnose some normality as abnormality in order to be certain."

This possibility of over-diagnosis produces a split amongst practitioners. Some see the machines' propensity to diagnose normal traces as abnormal as acceptable, others do not: A chief technician remarked that

> "It may give out abnormals when they are not. That is a step in the right direction. It is better for it to say 'query abnormal' when it is normal, than it saying its normal when its distinctly not normal ... so that someone has the chance to over-read it."

He saw the production of false abnormals as a small price to pay for a machine that can (theoretically) carry out the screening process. However, other practitioners offered an equally credible argument: When normal ECGs are diagnosed as abnormal, "inexperienced staff become alarmed", (registrar) "time and money is wasted on unnecessary referrals to hospitals, and patients become anxious" (consultant).

The question that arises from this discussion is, does it matter that the machines sometimes produce interpretations that are considered to be inaccurate and of little use? The human contribution is so evidently necessary and accepted, it is likely that the machine's mistakes do not matter because they will always be detected.

In practice the machine's normal decisions are not relied on by experienced operators in the largest hospital in the study. The ECG traces are routinely and automatically checked, and the experienced human experts do not accept the normal/abnormal distinction made by the machine, nor the full interpretations. Instead they rely on their own interpretations. The technician in charge of the pacemaker clinic in this department pointed out that patients attending the pacemaker clinic are routinely given an ECG, but in this situation,

> "When I get the ECG, I may read what it says on the top [the machine's analysis], but I don't take it in. I always interpret the ECG myself, because that's what we've always been taught here."

The users here seem to expect and accept the mistakes made by the machine. Rather than complain to manufacturers, they take on board the role of 'detector of mistakes'. Evidently the machine's performance is dependent on the charity of the humans who operate it. At this point it became clear that at this hospital, the largest in the study, many of the technicians never used the interpretation mode. They felt that checking the interpretation and recording another trace when the interpretation was unacceptable, was a waste of time. Furthermore, using the interpretation facility was in itself time consuming, because it involved asking the patient personal questions, typing this data into the machine and waiting for the machine to produce an analysis at the end of the recording. Very often the technicians either extracted the ECG from the machine before it was given a chance to print an interpretation, or else they requested a measurement-only report, and ignored the interpretation facility completely. Their justifications for this policy were wide ranging:

> "It takes so long to decide on it's diagnosis after it's printed the ECG."

> "It takes too long, and then it's sometimes wrong"

> "If you are busy you just can't wait - it takes ages. You could have done a whole ECG while you were waiting for it."

"Doctors here shouldn't need it - this is Cardiology"

"It is much too time consuming to punch in all the data that it asks for. Sometimes patients don't know how tall they are, let alone the drugs they are taking."

"I'm not confident to give it to them [doctors], in case I don't check it, and it turns out to be wrong, and they may take it at face value. Who would be at fault then?"

The technicians felt that their own interpretations were quicker and more accurate than those produced by the machine. So, in a large hospital such as this, no, it does not seem to matter that the machine is often unreliable. Human cardiological expertise is available literally 'on tap', and the machine's ECG analyses are over-read and often over-ruled.

But this is not the case everywhere. In the smaller hospital involved in the study, there were no resident medical cardiac staff and the technicians were more or less self-taught. There the machine was often relied on to give a definitive ECG interpretation. The sales representative admitted that,

"You will find that [some] Cardiology departments will use it as a definitive machine, to give them an interpretation ... I would say that of the market, one out of twenty buy it for that reason."

A doctor at this small hospital reported that

> "Most of the SHOs here are specialising in Opthalmology, and I'm sure they'd be the first to admit they are not crack ECG interpreters"

It seems that when these doctors require an interpretation of an ECG the machine's analysis is referred to. The technician in charge remembered that

> "We have had one or two people who have actually looked at the diagnosis and taken it as fact."

Bearing in mind the machine's mistakes described above, such reliance is alarming. The consequences of using an interpretive ECG machine are obviously different in this environment where human ECG interpretation skills are in short supply, than in the large teaching hospital. A visiting Consultant to the smaller hospital set out a vivid example of the problems associated with the use of the interpretive machine there:

> "When an abnormal ECG [according to the machine] comes up, the switch is flicked, and the patient is discharged. And if you put yourself in that patient's position, they've come to hospital, apprehensive, psychologically prepared to undergo an unpleasant experience. They get to the last stage and somebody flicks a switch, [on the basis of the automatic ECG interpretation], the trap door opens and out they go again."

In effect the machine was used as an exclusion device to decide who was and who was not fit to undergo anaesthetic and surgery.

5. SUMMARY

I have shown some of the problems involved with using interpretive ECG machines. In some departments, the machines are not used because experienced operators acts as a barrier - they may ignore the interpretation facility, believing they can do better themselves, or they repair the machine's shortcomings. This human barrier to their use is based on a recognition of the technological inadequacies of the machine. However, despite the practical problems highlighted here, interpretive ECG machines are becoming increasingly common in practice and continue to stimulate great interest. In establishments where human cardiac expertise is in short supply, staff look favourably on potential solutions to the skill shortage. Large and advanced medical units recognise other attractions of the machines - high quality presentation, glossy traces with no evidence of messy human intervention, and a series of useful amplitude and axis measurements. In both environments, manufacturers' 'special deals', which frequently offer interpretive machines at the price of non-interpretive models, are the final persuasion. One Chief Technician summed up the situation when he said,

> "Who wants to be the only Cardiology department without an interpretive machine?"

I have shown that both stages of the ECG interpretation procedure involve regular action rather than behaviour-specific action - the calcification of the normal/abnormal decision, and the classification of abnormal traces are situated achievements. The task of ECG analysis is complex, and lies towards the wide end of the venturi model, where the medical world meets the wider world. The wider world influences become evident when we consider the range of factors that influence all parts of the human achievement of ECG interpretation. It is understandable that the rule based functioning of the machine does not produce interpretations that always equate with human interpretations. Still the machines are popular in many departments, clinics and surgeries. Under these circumstances it is vital that the disadvantages and the advantages of the systems are made clear to all users, because the tendency is that in environments where human expertise is lacking,

> "People will believe anything that bloody machine writes, because it looks so formal, it looks so impressive. Their attitude is, this can't be wrong." (cardiology consultant).

Despite the inaccuracies in some of the interpretations, these machines blend into the social setting in the medical environments. I have shown that the human digitization, repair and general charitability toward the machine underlies this 'fitting in'. The human role in the machine's performance is essential. But in environments where human ECG expertise is in short supply, repair of the machine's inaccuracies **does not** always occur, and still the machines 'fit in'. It seems that staff who are unaware of the nature of competent human performance welcome the machine, as it appears - to them - to be an acceptable social prothesis. But it is not an acceptable social prothesis in the eyes of experienced practitioners. The reason for this is that it does not cross the boundary between ECG analysis (the discrete area in which it works) and the wider world - the arena in which experienced humans work.

It must be recognised that the machine demands the presence of an expert human interpreter. The danger of using the interpretive machines is that some physicians and technicians, who are not themselves expert ECG interpreters, are impressed by the nature of the trace and the mechanical interpretation. They are liable to accept the interpretation without considering the value of its content, or repairing it in the light of the particular patient under scrutiny. Only an experienced human can clear up ambiguities, explain anomalies to novices, check the validity of unusual measurement decisions, and decisively over-rule an interpretation. The role of the charitable, knowledgeable human is essential. In essence, these machines can be employed as tools, under experienced supervision. They are never suitable as replacements for experts, unless an informed decision has been made that the machine, with all the faults that have been highlighted here, is more acceptable than the available alternative. This decision is only likely in situations where the available alternative is no interpretation at all.

This case has involved looking at a machine which is attempting to do a task that involves a range of situated, regular human actions when people do it. The failures of the machine and the reasons why - despite these failures - it is still popular in Cardiology departments, have been discussed. The next case study is concerned with another machine - 'HELP' - that is tackling a wide range of human actions. Do the same explanations of charity, digitization, repair and human inability to recognise a competent or an incompetent performance, explain why the 'HELP' system is such a prominant aspect of health care at one particular American hospital?

FOOTNOTES TO CHAPTER 8

1. See Hartland (1993) for explanations and examples of various rhythm strips and abnormalities.

2. The machine has requested a repeat ECG because of the artefact present in some of the leads, not because of the incorrect limb lead application.

3. Transparencies of these traces are reproduced in appendix 1. When these are superimposed it is clear that although the traces are very similar, there are subtle differences. The doctors involved are able to ignore these differences.

4. See Hartland (1990b) for a more thorough discussion of the problems involved in 'Diagnosing the Normal'.

SECTION 2

CHAPTER 9

THE 'HELP' SYSTEM

CHAPTER 9 - THE 'HELP' SYSTEM

1. BACKGROUND AND HISTORY

The 'HELP' hospital information system differs from the other case studies presented: it is designed to tackle a wide range of tasks within the hospital, rather than a specific single measurement or diagnostic task. HELP has been under development at the Latter day Saints (LDS) hospital in Salt Lake City for more than 20 years. The LDS is a 520 bed private tertiary care hospital that acts as a major teaching centre for the University of Utah school of medicine. HELP facilities have gradually become part of the routine in most departments in the hospital. In their book, Kuperman et al (1991) describe HELP as a 'Dynamic Hospital Information System' - because it is continually being updated and improved, with new facilities and sub-systems being added.

HELP is the acronym for 'Health Evaluation through Logical Processing'. Throughout its development, the system has often been used as an example in the study of the use of computers in medical care. As a result, the evolution of the HELP system has been charted in great detail in the medical and research literature - Kuperman et al (1991) list 180 publications and 42 dissertations and doctoral theses that are based on HELP and its applications.

H R Warner is considered to be the main pioneer of the HELP system at the LDS. His work in the late 1950s involved the use of computers to make diagnoses on the basis of clinical information from congenital heart disease patients. These diagnoses were produced using Bayesian statistical analyses. This progressed in the early 1960s to work on computer processing of physiological signals. This marked the beginning of the field of patient monitoring. Successes here eventually led to the computerisation of many patient signals in the intensive care units (ICUs). As computer hardware and software improved, it became possible to directly interface laboratory blood gas analyses to a central computer at the LDS, so that laboratory data was automatically collected and transmitted to ICU, where it was used in the care of patients.

The next development at LDS was the application of ECG interpretation algorithms in 1968 - "The two first LDS hospital computer decision making applications to become clinically operational were interpretations of the ECG and blood gas results." (Kuperman et al, 1991, p. 9). ECG interpretation algorithms are accepted as being the forerunners of all other decision making and decision support systems offered in the current version of HELP. (See Warner, 1978) (1).

As more monitoring capabilities were introduced at the hospital, it became clear that integrating the outputs of these individual systems would be of greater value than the outputs of the individual systems themselves. In 1970 all clinical systems programs were combined to use a common data base, and this data base was the initial step in the formulation of the current HELP system.

There were four major objectives for HELP at this stage: 1. The system must allow for an ever-expanding medical data base. 2. It must possess medical decision logic, including criteria for diagnostic, therapeutic and alarm protocols. 3. It must serve the medical and administrative needs of the hospital. And finally, 4. the system must allow for effective research sub-systems that facilitate clinical research using the large data base in the system. (see Pryor et al, 1983). Since then, research into additional clinical applications of the system has continued and the number of decision support facilities has increased. A coding system for medical terms is now in place which has allowed expansion of the knowledge base holding the medical 'logic'. Throughout the 1970's the services offered by HELP were made available to more and more staff as terminals were installed on all the ICUs and some general nursing floors. The entire system was transferred to more modern computer hardware in 1979. This changeover incorporated facilities that allowed HELP to assist with the financial and administrative work in the hospital.

During the 1980s the decision making capabilities of HELP were expanded and more functions were added. In 1989 Gardner et al outlined the types of decision support that HELP was programmed to provide at that time. These were:

1. Alerting of staff to time critical or action oriented events such as abnormal laboratory values, vital signs or medical contra-indications.
2. Interpreting, by assimilating data results such as ECG morphology and rhythm, and interpretation of blood gas data.
3. Assisting through use of decision support to simplify actions, such as assisting with clinical orders.
4. Critiquing, which involves analysis and validation of decisions such as drug prescriptions.
5. Diagnosing by applying a medical 'model' for understanding the state of a physiological system - eg, diagnosing using Bayesian strategy.
6. Managing patients, using techniques such as the generation of protocols for ventilator control in critically ill ICU patients. (See Gardner et al, 1989, p. 96).

Add to this list the newer functions described by Kuperman et al (1991) - such things as surgery scheduling, Xray scheduling, all billing and administrative functions, ordering of supplies, nurse charting - and it becomes clear that "What started in 1967 as a computerised system to help monitor patients following open heart surgery has now dramatically matured." (Gardner et al 1992, p. 235). Now the daily operation and smooth running of the hospital depends on the HELP system.

2. THE COMPUTERISED MACHINE

Patient information is stored in HELP's clinical data base in two forms - a long term abstract of information to be used if a patient is readmitted to the hospital, and a short term comprehensive collection of all data gathered during the current admission. (Pryor et al, 1983; Burke et al, 1991). The patient reports and directives that the system produces are dependent on information from a wide variety of departments being entered into the integrated data base. Data entry is the responsibility of nurses, therapists, pharmacists and laboratory staff.

The system works on both a time drive and a data drive principle. In time drive mode the system activates at a specified time each day, or each week. In data drive mode the 'logic' is automatically activated whenever new patient data is entered. There is no need for manual activation of the system in either mode: HELP tirelessly monitors patient data, sifting through enormous quantities of clinical information.

The LDS hospital system is unique in the extent of the computerised services that it offers patients and staff. Kuperman et al point out that

"The research and development that have taken place at LDS hospital in physiological monitoring, systems integration, database development, man-machine interface and medical decision making makes the LDS hospital experience in the field of computerised hospital information systems unparalleled." (1991, p. 13).

In essence HELP is used for 1. data management, 2. information processing and 3. automated decision support. (Kuperman et al, 1991. p. 53). These writers explain that *Data* is the raw, uninterpreted elements used by decision makers - laboratory results, physiological signals, patient charges. The computer stores, organises and allows retrieval of this data. *Information*, by contrast, is a collection of data arranged in a manner that conveys meaning. An example is a display of various physiological readings that is presented in an organised way to a physician, which allows plans of action to be initiated. The distinction between data management and information processing can be blurred and somewhat arbitrary. *Automated decision support* is a different kind of computer task. For this the computer utilizes medical knowledge stored in the knowledge base and uses it to make patient-specific inferences.

Many instances of HELP using encoded medical knowledge for automated decision support are documented. The most comprehensive description of these can be found in Kuperman et al (1991). The infectious disease and drug alerting programs are two examples that fit into this category of automated decision support. The chief pharmacist explained that

> "the programs dealing with infectious diseases and drugs are used most often."

These two programs incorporate the following three capabilities:

1. Algorithms for determining which patients have hospital acquired infections, and for predicting which patients are likely to acquire these.
2. Guidelines for improving the timing of pre-operative antibiotic prescribing and for reducing the unnecessary use of post-operative antibiotic use.
3. Algorithms that alert the physician to the likelihood of an adverse drug reaction.

These three capabilities show the HELP computer undertaking 'human' tasks. The next section will concentrate on these three capabilities of the infectious disease and drug alerting programs, rather than on the data management and information processing capabilities.

3. THE FIELDWORK PART 1: THE INFECTIOUS DISEASE AND DRUG ALERTING PROGRAMS

The information in the following sections was collected in November 1991 during a visit to the LDS hospital. In order to attain understandings and native competence, I comprehensively trawled the literature about the HELP system, I had detailed conversations with designers and users of the system, attended demonstrations of new capabilities and observed parts of the system in routine use.

3.1 Detection of Hospital Acquired Infections

Between 5% and 8% of patients admitted to a US hospital develop a hospital acquired, or nosocomial, infection. The hospitals are required by the Joint Commission on the Accreditation of Health Care (JCAHC) to report all instances of nosocomial infections, explaining where and how the infection was acquired. The chief pharmacist explained how this is achieved at the LDS hospital:

> "In the infectious disease area, the computer diagnoses the nosocomial infections."

This facility has been available since 1984. The computer uses data from the microbiology laboratory and other branches of the HELP system, and 'knowledge' from the specific microbiology module of the medical knowledge base. The microbiology module has two components - the clinical microbiology laboratory system which produces routine microbiology results for patients, and the computerised infectious disease monitor (CIDM). The CIDM produces reports for the department of infectious diseases and for pharmacy. It is the CIDM that assesses every patient's data, accesses the knowledge base and produces a list of patients with nosocomial infections.

A CIDM report is automatically generated each morning for every patient. An alert is produced if the patient has a nosocomial infection (2). In addition to this, the CIDM has six other functions, and will alert the physician when any of the following are detected:

- An infection at a normally sterile body site.
- An infection due to a bacteria with unusual antibiotic sensitivity patterns.
- An infection in a patient who is receiving no antibiotics, or antibiotics which are inappropriate.

- A patient who is receiving an antibiotic that is not the least expensive.
- An infection that should be reported to the National Health Authority.

(As explained by the programmer).

Since the computer has been assigned the task of identifying nosocomial infections, the process has been speeded up, and the computerised method is generally considered more efficient than the manual method. The senior programmer explained that

> "The computerised surveillance identified 90% of the hospital acquired infections and the manual surveillance identified 76% [and] manual required about two thirds or three quarters of the infection control practitioner's time."

The accumulated information in the CIDM has also been analysed in retrospect. A statistical analysis of 24 variables from patients presenting nosocomial infections was carried out. The factors analysed ranged from sex, age, physician and disease, to type of surgery, use of urinary catheters and length of stay, etc. A predictive model was produced which allows the computer to pin-point the patients currently at risk of acquiring a nosocomial infection. The chief programmer explained how this works:

> "We run the patient through this model and the patient will get a score between 0 and 10. If your score is less than 1 you have got a 10.9% probability of getting an infection while you are in the hospital. But if your score is greater than 9 you have got a 97.6% chance. So the computer will tell us everyday which patients are at high risk."

About 35 nosocomial infection alerts are generated every day in the LDS hospital. (Evans et al, 1985). Evans suggests that identifying patients with nosocomial infections and isolating patients at risk of acquiring a nosocomial infection constitutes a "logical benefit". This he defines as "an improvement in the process of care, such as a reduction in time or medical personnel, that allows a process that is not efficient when performed manually to become efficient through the use of a computer." (Evans, 1991, p. 284).

3.2 Improving the Timing of Antibiotic Administration

In 1986 attempts were made to improve the timing of prophylactic antibiotic administration (3). These antibiotics should be administered 0-2 hours before surgery. A survey at the LDS in 1986 showed that only 40% of patients who should have received prophylactic antibiotics in this period actually did. (Evans, 1991). A computer program within HELP was designed to search all patients every day and identify those surgical patients requiring a pre-operative antibiotic (see Larson et al, 1989). The result is that,

> "Now there is a computer program out there that the nurses can actually run and it will tell them which patients should get a pre-operative antibiotic."

The administration rate during the 0-2 hour period has now increased to 95% of patients. Alongside this, significant decreases in the post operative wound infection rate have also been noted. The chief pharmacist explained the mechanism of this system:

> "Only 40% of people were getting their antibiotic within that time [0-2 hours pre-op]. And what we decided to do was use a reminder that was generated by the computer that would go and find all these people that needed an antibiotic and generate a sticker that would go in the chart so the surgeon would think about whether to give an antibiotic or not."

To make these predictions HELP uses information in the knowledge base which identifies procedures which require a pre-op antibiotic. Such procedures are apparently well recognised -

> "We can absolutely guarantee that there are certain surgeries that require antibiotics. There is literature world-wide to back you up on this one."

A further application of this program operates on a time-driven principle at 11am every morning. It identifies patients that have already undergone surgery and checks the time since the operation. If this is more than 48 hours and they are still receiving antibiotics, the program asks why. The policy decided on at LDS is that post-op antibiotics should stop after 48 hours, but this is sometimes overlooked. Assigning this task to the computer was designed to improve cost efficiency by stopping antibiotics on time. The pharmacist explained that after the computer has identified patients still on antibiotics after 48 hours:

> "It looks at the underlying diseases. In the case of thoracic surgery patients it checks to see has this patient had his arterial lines or swan ganz lines [out] because all thoracic surgeons want their patients to have antibiotics as long as they have their lines in. If

there is no evidence out there on the computerised medical record why the patient should still be on antibiotics the computer will generate an alert."

A report in 1990 concluded that computer surveillance in this field is an efficient and promising means of identifying errors in microbial prescribing. (See Pestotnik et al, 1990).

The continuous antibiotic monitoring of all patients that HELP performs is facilitated by the modular knowledge base and integrated patient data base. Burke et al's paper (1991) offers a detailed description of these processes.

3.3 Detection of Adverse Drug Events and Reactions

As many as 30% of patients may experience an adverse drug event (ADE) when in an American hospital. 0.31% of patients suffer an ADE that is fatal - this amounts to between 60,000 and 140,000 patients every year (Classen et al, 1991).

ADEs of two types arise. Type A is a predictable event - a drug/drug, drug/food or drug/laboratory interaction. Type B reactions occur either when patients have an allergy to a drug - which is not predictable unless the patient is aware of the allergy - or when the normally prescribed dose of a drug is too high for the particular patient under observation.

Hospitals are required to report all ADEs to the JCAHC and this was traditionally done manually by the nurses. The nurses "had to describe the reaction and what was done about it, decide what drug caused the reaction, and have the nursing supervisor and prescribing physician sign the report." (Evans, 1991, p. 286)

Relying on this voluntary reporting is potentially cumbersome and ineffective (Classen et al, 1991), and in the year May 1988-April 1989, only nine ADEs were reported by this manual method at the LDS.

The knowledge engineers and computer programmers set about designing a program that would do the job of checking for ADEs automatically. The resulting module is driven by patient data. All patient drug data, laboratory data, food data, and known allergy data is checked every day by the program. This system is easy to use, and far less time-consuming than the manual method. The programmer explained that,

> "They [nurses] just enter the data. Every time that gets entered the information will activate the data driver. And the data driver activates the knowledge base."

The system has been in operation since 1989. In the first year, 401 ADEs were identified and reported. The following year, with an extension of the knowledge base, this increased to 597. By 1991, Classen et al (1991) reported that the figure was 700. This is an increase of almost 80 fold on the manual figure of 9 cases per year reported.

The computer has access to every aspect of the patient data, and is able to constantly monitor all the variables, so that all ADEs for every patient are pin-pointed. Of the ADEs detected by HELP, 95% were classified as moderate or severe, and resulted in a change in therapy or duration of hospital stay. Recognising ADEs quickly allows for early cessation of the causative agent and potential prevention of more serious manifestations, which could have resulted in a longer period in the hospital. Clearly the ADE monitor is another example of what Evans (1991) terms a 'logical benefit'.

The program has been improved in an attempt to reduce the number of false positives it produces (4). New 'logic' has been added that allows HELP to check for other patient specific reasons for an unusual result, before suggesting an ADE. The knowledge engineer explained this concept:

> "For example, if the patient has a creatinine clearance of less than 50ml/min, that will activate the knowledge base as a possible adverse drug event. But then the knowledge base goes to check, say, does this patient have any other reason why he is having other renal problems? It checks the underlying diseases - is he scheduled to have kidney transplant or whatever, so it checks the underlying diseases and surgery information. If it cannot determine why the patient would have a creatinine clearance that low other than an ADE, it generates an alert."

In a similar way to which the HAI program has been extended to predict the patients at most risk of acquiring a nosocomial infection, this program is being developed to predict which patients will be likely to suffer from an adverse drug event. A retrospective analysis of patient variables is being conducted so that new patients at risk can be isolated.

4. DISCUSSION

Neither the infectious disease program nor the drug alerting program are foolproof - they may suggest nosocomial infections or ADEs in patients where they are not evident; they may suggest that post-operative antibiotics should be stopped, when the patient under scrutiny actually requires an extension to the usual 48 hour post-op administration. The programmers have two methods of reducing these false positive alerts: First by continuously updating and improving the knowledge base, adding new logic so that the computer will consider as many alternatives as possible before issuing an alert. Secondly, by attempting to educate the doctors to inform the system when a patient is somehow unusual in their presentation and requires drugs or treatment for longer than usual guidelines.

Doctors at the LDS are reluctant to become involved with inputting data to HELP, so patients who are somehow unusual, and exceptions to the rules and logic in the knowledge base continue to crop up, and alerts continue to be generated inappropriately. In these cases it is then up to the doctor to over-ride the computer's suggestions. The chief pharmacist pointed out that

> "You are still going to have to rely on your personal experience to say 'do I believe this and does it apply in this particular patient?'. Because the computer cannot know everything about every patient."

Although they can make mistakes, the three programs that I have selected and described all offer 'logical benefits' (Evans 1991, Table 2).

They all allow processes that were not efficient when performed by humans to become efficient through the use of the computer: Identifying hospital acquired infections using the computer is now more efficient and takes less time than it took manually; the administration rate of prophylactic antibiotics during the optimal period has improved from 40% to 95% of patients; identifying and reporting ADEs has increased from 9 cases a year by manual methods to 700 cases by computer. The computer is doing these jobs far more efficiently than the humans did. How does the computer achieve such dramatic increases in efficiency?

The chief pharmacist suggested that an important factor is the nature of the area of drugs and infectious disease. He explained that,

> "Infectious diseases and pharmacy is a field that many physicians, nurses and even pharmacists cannot keep abreast of everything. There are so many factors, so we started to look at the areas in infectious diseases and pharmacy where we could make the greatest impact with the computer."

The areas that these programs cover are broad and the data that is amassed each day for over 500 patients is vast. Systematic computer searching is a more effective means of dredging the data, than are manual searches. In order for HELP to produce lists of patients with nosocomial infections, patients with ADEs or patients requiring pre-op antibiotics, it must have access to all the information. The integrated data base in HELP makes this feasible. As the programmer explained,

> "The knowledge base doesn't just activate the data from the pharmacy, it has access to everything else, surgery data, previous data, X ray data - everything."

The computer is able to scan this data constantly. This is beyond the capabilities of the human staff, firstly because people become tired and bored when asked to search routinely through large amounts of data, much of which is acceptable data. The programmer emphasised this point, saying

> "If the human is checking a thousand patients and is only going to find a situation in two of those, there is the chance that the human mind wouldn't even pick that up because it is so tedious. But with the computer it doesn't matter. It can do a thousand patients every day."

Beyond this, time constraints mean that

> "A human cannot look and monitor everything, every test result. [But] The computer is able to constantly monitor everything."

The implication is that the systematic, rule-based searches performed by the computer are an effective way to do jobs that humans cannot, or choose not to, perform. The result is that the computer equips the medical staff with more information and data than they would otherwise have access to. When the computer-produced information becomes available, decisions about what to do next are made by the medical staff. At this point the human staff take over from the computer. A knowledge engineer in the radiology unit explained the necessity for the human involvement after the computer provides data and makes suggestions:

> "The judgement phase, the common sense phase happens after the computer makes its suggestion. The computer makes a suggestion that is filtered through the human who works with all the imponderables that a computer can't ponder."

At this level of operation the computer is a provider of information rather than a decision maker. This view is reflected in the published literature on HELP: For example, Evans (1991) states that "The computer can provide timely and important information but it is the human user who must apply the information." (p. 287). The knowledge engineer stressed this point:

> "Somebody can then take that information [provided by HELP] and they then make the final decision or choice. But the computer is never going to replace the physician. The computer can help you and recommend and the computer can alert and can suggest. It is the doctor who really has the more broader view of everything, and he makes the ultimate choice. This [computer] just provides as much information as possible by doing these things. He [the doctor] can't do all this stuff."

This is the crux of the issue. The computer does things that the doctor does not do. (See also Gardner et al, 1989, p. 97). The object of these HELP modules, which use medical knowledge, is to enhance processes that the humans themselves are unable to do, or do badly. The knowledge engineer agreed with this interpretation, and explained his view as follows:

> "We are constantly thinking, and this is what I do, I design and think about things ... this isn't working manually, things that the computer can do that you cannot do manually or something we can do manually but the computer can do better and faster."

ADE monitoring and prophylactic antibiotic screening are examples of tasks that the medical staff can do, but usually do badly. So too is identifying nosocomial infections - in theory it can be done manually but would be far too time consuming. Similarly with identifying patients taking post-op antibiotics for too long. The pharmacist felt that,

> "You could do that manually if you hired five people just to go through everybody's record every day."

These computer functions are not 'replacing' humans in medical tasks. Warner (1980) stressed this point. He stated that the aim of the system is "to look for ways of handling things that are presently difficult for a human to handle." (p. 77). The computer's success in these areas rests on the nature of the computer's abilities in relation to human abilities. Computers are able to sort through large volumes of patient data, day in, day out, without becoming bored or careless. They follow pre-programmed rules, applying them to all cases, to isolate individual cases. The system does not respond to patient specific variations, but follows rules in the same way every time, carrying out a behaviour-specific response. Its tasks exist inside microworlds at the narrow point of the venturi - no external factors affect the searches and the mathematical calculations. Responding to individual patient's idiosyncrasies is left to the medical staff who oversee and evaluate the computer's output in the light of real world circumstances. In this way, any instances where the decisions made by HELP do not fit the particular patient under scrutiny are identified. The patient specific knowledge and context based decision making performed by the staff allows the HELP decisions to be used practically. So, the computer performs a range of tasks in the way that a human would if they were using behaviour-specific action, and the humans check these responses using regular action. The staff are thus provided with useful information that allows them to proceed with the next stages of patient management more effectively. As Evans (1991) points out, "The appropriate use of the information provided by the computer makes the application successful." (p. 287).

Studying the chapters in Kuperman et al's book (1991) where the capabilities of the HELP system are described in detail, almost all fall under this umbrella - they are modules programmed to perform tasks that the medical staff are ill equiped to do. As a result the computer provides the staff with useful data. This applies to the laboratory results that are displayed at various sites all over the hospital, the billing facility that adds up patient costs automatically as the patient progresses through their stay, surgery scheduling, ordering facilities, etc.

Each of these programs is a useful addition to the running of the hospital, prompting staff and providing information. The staff role is one of an expert overseer of the computer output. Together the computer and the staff perform well, but the computer is not replacing a human making medical decisions. However, the programmers and developers at the LDS have recently branched out in this field, developing 'protocols' to direct the management of intensive care respiratory patients. The computer has been programmed with the protocol instructions which it follows and from which it produces directions for the management of patients. One of the knowledge engineers explained the significance of this progression, in relation to the other tasks that HELP was programmed to do:

> "You come up with any of those alerts [such as those described above] and ask a doctor if he know it, and he will say 'sure I knew it, I just happened to miss it at that particular point in time'. But the things we are doing now on ICU with the respiratory system, they are a step beyond that."

The next section looks more closely at this 'step beyond'.

5. THE FIELDWORK PART 2
COMPUTERISED CLINICAL PROTOCOLS - A STEP BEYOND?

5.1 Introduction

Engineers at the LDS have installed protocol programs into the HELP computer. These are designed to enable HELP to direct the ventilator therapy of patients with Adult Respiratory Distress Syndrome (ARDS). Using protocols demands that doctors forfeit their own individual methods of treatment in favour of computerised 'consensus' based (protocol) decisions. In this capacity, HELP seems to be taking over a human task involving complicated decision making processes.

ARDS is an acute respiratory failure condition requiring artificial ventilation. "There is no uniform and universally accepted definition of the syndrome" (Morris et al 1989, p. 138), although common elements are arterial hypoxaemia (5), radiographic infiltrates (6), and clinical expression of respiratory distress (7). There are an estimated 150,000 cases of ARDS annually in the United States, and patients who survive do generally resume productive lives. In 1979 a subset of ARDS patients meeting pre-specified ECMO criteria (8) were singled out and found to have a survival rate of about 10%. By contrast, in 1984 Gattinoni et al published results of a new therapeutic program for ARDS patients satisfying the ECMO criteria, which appeared to achieve a 77% survival rate.

The new treatment used involved a 3-step therapy program: First, pressure controlled inverse ratio ventilation (PCIRV), and then either continuous positive airway pressure (CPAP) if the patient improved on PCIRV, or low frequency positive pressure ventilation with extracorporeal CO2 removal (LFPPV-ECCO2R) if the patient failed to improve on PCIRV. The LFPPV-ECCO2R is designed to underinflate the lung and limit the ventilatory rate, so as not to further damage an injured ARDS lung (Morris et al 1990).

This vast increase in survival rate prompted international interest, despite there being no control or randomization during the trial. Researchers at the LDS were anxious to repeat the trial under more regimented conditions. The proposed study at the LDS involved assigning ARDS patients satisfying ECMO criteria to two groups. Traditional treatment - continuous positive pressure ventilation (CPPV) - was to be administered to the control group and the new therapy was to be administered to the other group. A problem with this simplistic design was feared, which the researcher at the centre of the project explained:

> "We may create an environment where the patients who were having the new therapy would have all kinds of attention to detail

> and senior people looking in on them. Whereas the ones that were randomised to the control would get the medical student and the student nurse."

It is recognised in the literature that "variations in patient care can be seen not only among hospitals but even among ICUs within a single hospital, and among physicians within an ICU." (Morris et al, 1990, p. 229), so it is possible that the different groups in this trial could have received different levels of care. With this in mind the researchers attempted to ensure that both groups received the same quality of care so as to improve the credibility of the trial result (Morris 1992). There were four major aims: (1) that uniform care be given to both groups, with (2) equal intensity of monitoring and (3) equal frequency of monitoring, and finally, (4) consistent decision making logic be used in both groups. (Morris et al 1989, Morris et al 1990, East et al 1990, Morris 1992). Specific treatment protocols were conceived in an attempt to ensure that these criteria were met, and to direct the respiratory management of all the patients with ARDS in the trial, whether they fell into the experimental or the control group. One researcher suggested that equalising quality and quantity of care for both branches would also make the study "far more scientific".

5.2 Protocol Development

East et al (1990) state that "Physicians' decisions are influenced by many different things including: the journal article that they read yesterday; the research meeting they attended; the colleague they spoke to last week; the last drug representative that visited and the current patient that they just spent the whole night trying to save." (p. 566). Clearly, physicians make decisions in different ways, depending on these factors. If protocols are to be created to direct the execution of any task, all involved parties have to forfeit their own preferred method in favour of the decision given by the computer. A six stage method was used at the LDS to design the ARDS protocols. It involved discussion sessions with various personnel, production and testing of paper based protocols, computerisation of the protocol logic, validation and eventual release for routine use. A diagram representing this process is shown below - figure 9.1. (See also East et al 1990; East et al 1992).

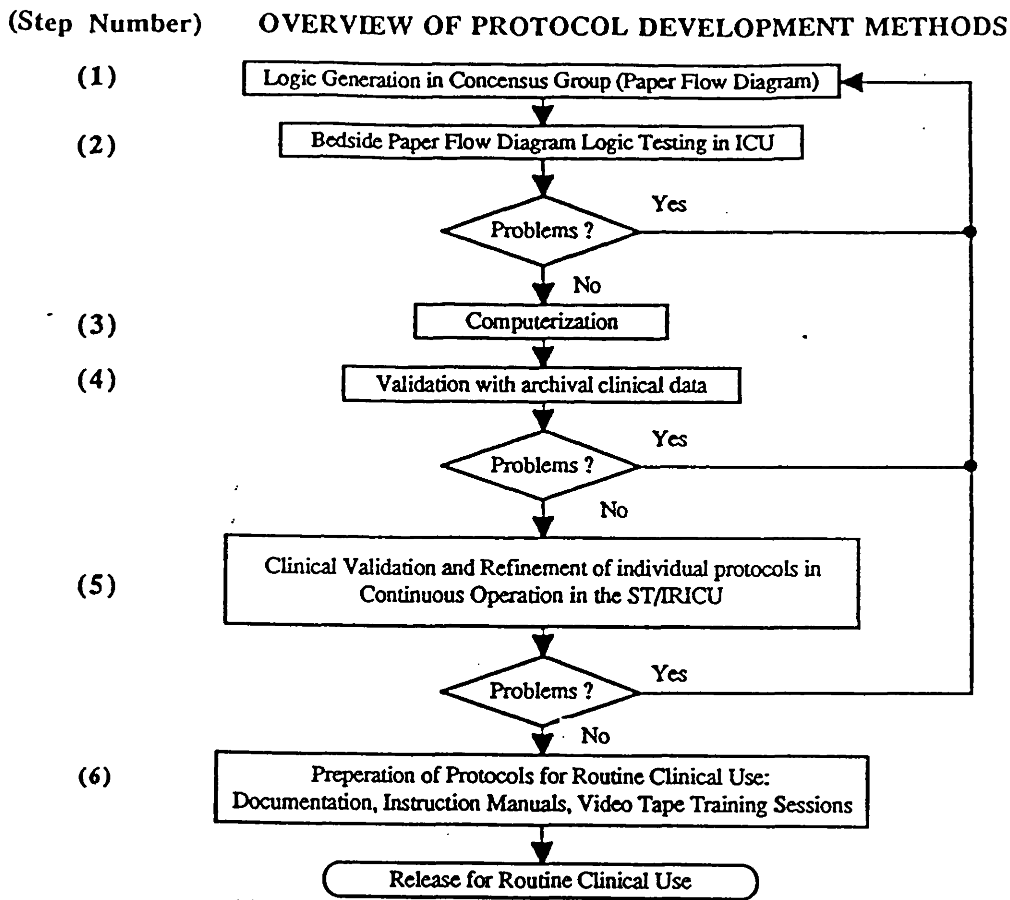


Figure 1: Overview of current protocol development process

Figure 9.1

A researcher extensively involved with organising the production and perfection of the protocol logic and its introduction into practical environments, explained that initiating discussion and forming consensus amongst the physicians had not been easy. The argument she had used to help persuade them to collaborate went as follows:

> "Look, everybody has a different way [of managing the ARDS patients] and still we get a terrible outcome (9), so what is the problem with saying 'look, I don't know if my way is better, but we need to pick a [specific] way, all do it, and then see if it has any effect on the outcome'."

She admitted that the process had been long and tricky, involving a great many meetings and discussions, and that early on in the process she doubted whether the project was feasible:

> "During the first five hundred hours I thought 'we are not going to be able to do it'. "

The doctor in charge of the ICU echoed this, explaining that developing this first set of protocols had been

> "bloody and brutal, because getting physicians to give up their own stylistic ways of doing it is very very hard."

He pointed out that

> "If you look at the decisions that are made in medicine, what you are going to find out is that at least 50% of what we do, we don't know if it is right or wrong. And what it came down to was saying 'this [consensual method] may not be our way, but it is a way, it is a way that we can accept and it is not different from letting the residents make decisions.' And we do that all the time."

After much debate, tests and trials, the protocols were completed, computerised and ready for application at the bedside.

The protocols respond to changes in arterial pO2 levels, suggesting an increase or decrease in therapy, or giving instructions to wait and withhold any changes in therapy. The protocols were intended to be fully operational by the time the trial comparing the new therapy for ARDS with the traditional therapy was started. Figure 9.2 (overleaf) shows a portion of protocol instructions for a patient on ICU at the LDS on 31 January - 1 February 1989. (This chart reads from the bottom upwards).

The notion of using computer-generated decisions to direct therapy is important. Have the human professionals handed over control of this aspect of care to the computer? To what extent, if at all, is human involvement still required? If the computer's decisions are acceptable to the physicians in all cases, it is a major achievement for both the knowledge engineers and the programmers. How well do the protocol instructions 'perform' and how does their utilization fit into the regime of the ICU? Researchers at the LDS have published the results of the clinical trials of the new ARDS treatment. These results fostered increased interest in the protocols themselves, as explained below:

PATIENT #114666

DC 11 FC 36 TIME 02/01/89.08:42
REDUCE FiO2 BY 10 % DRAW ABG IN 15 MIN

DC 11 FC 36 TIME 02/01/89.07:38
REDUCE FiO2 BY 10 % DRAW ABG IN 15 MIN

DC 11 FC 36 TIME 02/01/89.05:28
CONTINUE TO MONITOR PATIENT FOR 2 HOURS AND THEN DRAW ARTERIAL BLOOD GAS

DC 11 FC 36 TIME 02/01/89.04:16
ACCEPTABLE OXYGENATION 4 HOUR WAIT WITH 1 HOURS 9 MINUTES REMAINING. DRAW ARTERIAL GAS AT 5:25 HOURS

DC 11 FC 36 TIME 02/01/89.03:11
CONTINUE TO MONITOR PATIENT FOR 2 HOURS AND THEN DRAW ARTERIAL BLOOD GAS

DC 11 FC 36 TIME 02/01/89.02:35
CONTINUE TO MONITOR PATIENT FOR 2 HOURS AND THEN DRAW ARTERIAL BLOOD GAS

DC 11 FC 36 TIME 02/01/89.01:32
CONTINUE TO MONITOR PATIENT FOR 2 HOURS AND THEN DRAW ARTERIAL BLOOD GAS

DC 11 FC 36 TIME 02/01/89.00:18
ALTERNATING MODE 4 HOUR WAIT WITH 1 HOURS 2 MINUTES REMAINING. DRAW ARTERIAL GAS AT 1:20 HOURS

DC 11 FC 36 TIME 01/31/89.23:17
ALTERNATING MODE 4 HOUR WAIT WITH 2 HOURS 3 MINUTES REMAINING. DRAW ARTERIAL GAS AT 1:17 HOURS

DC 11 FC 36 TIME 01/31/89.21:23
INCREASE FiO2 BY 10 % DRAW ABG IN 15 MIN

DC 11 FC 36 TIME 01/31/89.20:36
CONTINUE TO MONITOR PATIENT FOR 2 HOURS AND THEN DRAW ARTERIAL BLOOD GAS

DC 11 FC 36 TIME 01/31/89.20:33
REDUCE FiO2 BY 10 % DRAW ABG IN 15 MIN

DC 11 FC 36 TIME 01/31/89.20:33
REDUCE FiO2 BY 10 % DRAW ABG IN 15 MIN

DC 11 FC 36 TIME 01/31/89.17:21
CONTINUE TO MONITOR PATIENT FOR 2 HOURS AND THEN DRAW ARTERIAL BLOOD GAS

DC 11 FC 36 TIME 01/31/89.13:21
CONTINUE TO MONITOR PATIENT FOR 2 HOURS AND THEN DRAW ARTERIAL BLOOD GAS

DC 11 FC 36 TIME 01/31/89.12:46
CONTINUE TO MONITOR PATIENT FOR 49 MINUTES AND DRAW ARTERIAL BLOOD GAS AT 13:34 HOURS

DC 11 FC 36 TIME 01/31/89.12:02
CONTINUE TO MONITOR PATIENT FOR 1 HOURS 32 MINUTES AND DRAW ARTERIAL BLOOD GAS AT 13:34 HOURS

DC 11 FC 36 TIME 01/31/89.10:38
CHANGE MODE OF VENTILATION TO A/C RATE/MIN EQUALS 42 DRAW ABG IN 15 MIN

DC 11 FC 36 TIME 01/31/89.08:00
CONTINUE TO MONITOR PATIENT FOR 1 HOURS 50 MINUTES AND DRAW ARTERIAL BLOOD GAS AT 9:50 HOURS

DC 11 FC 36 TIME 01/31/89.07:06
ACCEPTABLE OXYGENATION 4 HOUR WAIT WITH 2 HOURS 31 MINUTES REMAINING. DRAW ARTERIAL GAS AT 9: 6 HOURS

DC 11 FC 36 TIME 01/31/89.05:20

Figure 9.2

fig 9.2

5.3 The Clinical Trial

The exact dates and details of the clinical trial are not immediately clear from the literature. For example, in 1989 Morris et al forecast that 60 patients would be included in the trial over a two year period and that "stratified randomization with blocking will be used and all personnel will be blinded" (p. 143). Stratified randomization was to be limited to age (<40 years / >40 years), and cause of ARDS (trauma / non trauma). Whereas in 1990, Morris et al describe a "randomized, prospective single centre clinical trial" (p 228). 40 randomized patients stratified by age and by the presence or absence of trauma were referred to in this paper, and it is stated that "Blinded randomization with blocking is being used." (p. 229). Furthermore, the papers by Morris (1992) and Morris et al (1992) both refer to a "prospective randomized clinical trial", of 40 patients satisfying ECMO criteria for ARDS, which took place between August 1987 and April 1991. The 40 patients included were divided - 19 in the control group and 21 in the new treatment group. Overall survival was reported at 38% (42% in the control group and 33% in the new treatment group). This survival rate represents almost a four-fold increase in the traditional survival rate.

These descriptions of the trial(s) do not make clear whether the patients were blind to which group they were in or whether the staff were blind to this detail. Or were both staff and patients blinded? Who was aware of which patients were in which group? What exactly 'prospective' means in this context is not explained. Are the 40 patients referred to by Morris and Morris et al a subset of the 60 patients reported on in the 1989 paper, or are they a different set of patients?

Wallace (1993) has explained that these papers, with inconsistent details of the ARDS treatment study, all refer to the same trial, which took place between August 1987 and April 1991. The final results of this trial are reported in Morris et al (1992). The conclusion that the researchers drew from the results is that survival of the ARDS patients increased when the protocols were used to direct treatment, to a level around 35%. This was the case for the control group who received the traditional therapy and for the experimental group receiving the new treatment. On this basis they suggest that extracorporeal CO2 removal therapy is not recommended for treating ARDS - there is nothing to suggest that the therapy itself improves survival. The survival rate of both groups in the trial was a surprise and researchers felt that the increase was due to the use of strategic protocols to direct care. The focus of attention suddenly shifted from the usefulness of the new ARDS therapy to the perfection of the protocols. The clinical director of the ICUs at the LDS summed up the intended direction of the department:

> "If we can find the funding, [we are] going to launch into a lot of different protocols, for doing standardised care, because we are totally convinced that standardised care is better care. You really have got better outcomes."

He felt that the standardising of care was the essence of improving care - it allows everyone in the team to be completely familiar with the processes and procedures, and to work in a routine manner. His opinion is that:

> "The standardisation process in itself is very good. And I think that it improves care a lot. And I think that the protocols showed us that."

This notion of standardised, algorithmic rules to control an essential aspect of a patient's therapy portrays an extremely 'scientific' image of medicine. It runs contrary to any notion of medical decision making being context related or intuitive. As such, the idea of protocol directed care offers a stiff challenge to those who see medical decision making as a situated, non behaviour-specific action. Closer attention to the way protocols are used on the ICU, and how they fit into the regime of the environment is necessary. What contribution do staff make to the apparent usefulness of the protocols?

5.4 The Protocols In Use

The practical use of protocols on the ICU involves the computer producing patient-specific instructions based on arterial oxygen levels. The system is data driven, activated by patient input regarding the arterial O2. The instructions generated by the computer are not implemented by the computer - they are assessed by the staff, who choose whether or not to follow the directive.

When the protocols were under development and first introduced at the bedside, staff followed about 60% of the instructions. The other instructions were ignored or challenged. As the protocols were improved, decision rules adjusted and logic updated, the percentage of the time that they were followed increased to nearer 90%. (Henderson et al, 1990; Henderson et al 1992; East et al 1992). The system as it is now leaves "the physician free to decline to follow a protocol instruction, if he or she has a defensible reason." (Morris, in press, p. 8). These challenges to the system are recorded and discussed, so that their validity can be assessed. When the challenges are deemed 'valid', appropriate changes are incorporated into the system in an effort to perfect the program.

The figures for staff compliance suggest that staff are - on the whole - following the protocol instructions. In effect they have delegated this job to the machine for about 90% of the time. But compliance is not maintained at a constant level: Drops in the level of compliance with the protocol instructions by the staff occur periodically. These drops are attributed to "the introduction of new logic, rotation of new clinical staff into the ICU and identification of previously unencountered clinical problems." (Henderson et al, 1992, p. 277).

A paper by Henderson et al (1992) is concerned with the proportion of the instructions generated by the computer that are considered 'correct'. A 'correct' instruction is defined as one that is the same as was specified in the original paper-based version of the protocols. The results showed that 90.2% of the instructions were considered - in retrospect - to be 'correct', with less than 10% 'incorrect'. In a different paper (1990) Henderson et al reported that 11.4% of the instructions generated by the computer were, in retrospect, 'incorrect'. Almost a third of these 'incorrect' instructions were implemented, which suggests that staff are sometimes inclined to follow instructions, regardless of the appropriateness of that instruction. The authors claim that in retrospect, a reason why an incorrect instruction was generated could be found in all but 1.5% of instances, and that this number of incorrect instructions will diminish continually as the program is improved. It is clear that

(1) medical staff do not comply with the instructions all the time - about 10% of the instruction are challenged or over-ruled.

(2) The computer does not always produce 'correct' instructions, as were decided on and laid down in the original version of the paper protocols - about 10% of the instructions are, in retrospect, labelled 'incorrect'. These are not exactly the same 10% as were challenged, because a third of these incorrect instructions are implemented.

(3) The protocols are only in operation for 86% of the day (Morris et al 1992, p. 85). The rest of the time the patients are taken out of protocol, because the circumstances are unusual or the physiological feedback does not represent the patients 'real' condition.

These three points should be borne in mind when attempting an analysis of the machine's protocol capabilities and it's apparent performance of a human activity. If the patient is under protocol control for 86% of the day, only 88% of the protocol instructions generated during this period are followed (Morris et al, 1992), and approximately 10% of the instructions are incorrect, then human directed treatment is clearly being administered some of the time. In the next section I examine the use of the protocols, with particular attention to the three factors outlined above. This analysis shows the essential nature of the human contribution to the care of the ARDS patients.

5.5. Closer Examination Of The Protocols

5.5i Staff Non-Compliance

Staff non-compliance occurs in about 1 in every 10 instructions. Reasons for non-compliance can often be formulated in retrospect by the designers, but this does not mean that instances of non-compliance can be anticipated in advance. Alterations to the program can only be made after staff have rejected a particular directive.

Staff knowledge of what is and what is not an acceptable treatment decision is essential if challenges to the system are to be made, and inappropriate directives ignored. Henderson et al state that one of the causes of a drop in compliance with directives is the occurrence of previously unencountered clinical problems. When a patient's condition varies from the expected pattern, and the patient shows individualised characteristics, the protocol produces directives that staff may not comply with. Protocols are not designed to respond to unusual, unexpected clinical problems in the way that a human would and the role of a human 'overseer' in these situations is essential.

5.5ii 'Incorrect Instructions'?

In a retrospective analysis, a portion of the directives issued by HELP are labelled 'incorrect'. Some of these are followed by the staff, some are rejected. There are two problems with 'incorrect' instructions: First, although no fatalities have yet occurred as a result of an incorrect instruction being implemented (Henderson et al 1990), this cannot be ruled out in the future. Second, and more fundamentally, why are incorrect instructions produced? It may be because the patient is in a chaotic or emergency situation which the protocol does not cover. Or it may be that the particular patient under scrutiny is for some reason an exception to the rules which the protocol proceeds by. In either case the experienced doctors' skill is necessary to direct care in a way that is acceptable for the particular patient and the particular scenario.

5.5iii Time That Patients are Out of Protocol

The protocols were designed only for management of the stable critically ill patients. The patients are removed from the protocol environment when they enter unstable critical phases, or when they exhibit unusual characteristics. For these acute or emergency situations protocols are abandoned and doctors' traditional skills and ability to react intuitively to unexpected developments are called on so that the patient is managed successfully, through regular action. As the chief researcher pointed out,

> "If the patient was in a crisis and had to have some acute problem taken care of, these scenarios - we didn't write the protocols for those crisis situations, we wrote the protocols for a patient who was in a stable, albeit very sick state of affairs."

6. DISCUSSION

These three issues have a common thread linking them - the notion that the protocols are designed only for treating stable critical patients. When instability is introduced into the system, directives are ignored, incorrect directives may be produced, or ultimately the patient may be removed from the protocol environment.

The protocols are designed to work when the patient's responses fit into a pre-defined set, for which the optimum management has been anticipated and programmed. Within the boundaries that protocols operate, the application of specific management strategies for specific physiological changes seems to work well (10), possibly increasing the chances of survival to about 35%. In this respect it is more successful than the mix of responses to a poorly understood condition that a range of ICU doctors would produce. As shown above, the doctors use regular action. The protocols are an attempt to replace this regular action with a behaviour-specific response. As Morris and Gardner point out "the protocols provide a standard therapeutic response to the arterial hypoxaemia in mechanically ventilated patients with severe ARDS." (1992, p. 510). This does not mimic the regular human action, but replaces it with something which, in this instance appears to be more effective in many cases.

It is important to remember that the protocols are designed to respond to pO2 levels - this is a small part of the overall care of an unconscious ICU patient. It is a small task, located within the broader context overall patient care, but isolated from it. The computer is operating on a very narrow set of variables. The computer uses a behaviour-specific response and produces acceptable results within this narrow field. However, when the patient's responses move out of this microworld situation, towards the wide end of the venturi, a doctor is needed to care for the patient using regular action to respond to unexpected events.

Areas where behaviour-specific responses are acceptable are difficult to locate in the medical world. This particular area exists because regulation of pO_2 is intrinsically isolated and because the responses of unconscious ICU patients are more predictable than those of fully conscious general medical patients. The staff involvement in the management of ARDS patients occurs when the patient's signs or responses become too 'unusual' for the computer. When broader factors begin to influence a patient's condition - for example, when a patient begins to regain consciousness, or when the physiological information falls outside the boundary of the computer's competence - the doctor steps in. The machine is by-passed and knowledge from 'beyond the microworld' is used to manage the patient.

7. SUMMARY

The protocols are an exciting development on the ICU. They respond in a pre-determined fashion to changes in arterial oxygenation levels. Their behaviour-specific response works in a manner that staff find acceptable for much of the time. The role of the experienced doctor or respiratory therapist remains essential for overseeing the system, because the protocols only deal with stable patients who are confined to a 'microworld' of medicine at the narrow part of the venturi. Human intervention is needed when these responses develop into the less predictable undigitized responses of an unstable patient nearer the wider medical world. Human expertise is essential for carrying out the aspects of the job where regular action and situated responses to unexpected circumstances are needed. Protocol directives serve only as advice and do not directly influence the treatment of the patient because protocols utilise behaviour-specific responses. The human intermediary is still an essential link in the treatment of ARDS patients. The protocols are a step beyond the other HELP programs, but are evidently less of a 'Step

Beyond' than seemed to be the case at first. They have not crossed the boundary between behaviour-specific action and regular action. Neither have they stepped into the area where medicine meets the real world. Protocols operate in a microworld where behaviour-specific action suffices.

FOOTNOTES TO CHAPTER 9

1. Warner (1978) says that "Electrocardiography has been the proving ground for many of the basic concepts of medical decision-making by computer. This paper describes a system [HELP] in which these concepts have been generalized for application to the whole field of medical practice." (p. 115)

2. This is an example of the system operating in time driven mode.

3. Prophylactic is a term used in medicine which means to treat in advance, or prevent before symptoms occur.

4. A false positive occurs when the computer suggests an ADE although none is present.

5. This is deficient oxygenation of the arterial blood.

6. These show as shadows over the lung on an Xray, whereas chest Xrays normally show the lung without shadows.

7. The clinical signs of respiratory distress are dyspnea - laboured and inadequate breathing, and tachypnea which is rapid but shallow breathing. Both lead to inadequate oxygen levels in the blood. Thanks to Dr. Ann Miller at the Respiratory Department, Royal United Hospital, Bath, for her explanation of these terms.

8. Extracorporeal membrane oxygenation - ECMO criteria. See Morris et al 1989, Table 2.

9. A terrible outcome in so far as the average survival of ARDS patients is traditionally about 9%.

10. This is a broad assumption based on the ARDS trials at the LDS hospital, and it is made bearing in mind the lack of controls used during the trial.

SECTION 3

CHAPTER 10

SUMMARY, CONCLUSIONS AND SUGGESTIONS

CHAPTER 10 - SUMMARY, CONCLUSIONS AND SUGGESTIONS

1. INTRODUCTION

There is a lack of practical assessments of medical computers in everyday environments. Contradictory images permeate the literature on artificial intelligence in medicine, but most of this work is based on theoretical analysis or controlled clinical trials of computers in research environments. By contrast, the investigation presented here was conducted using a participatory methodology to look at computers being used in working medical environments.

In view of the importance of the issues under discussion, a sociological investigation of this area is legitimate, and because medicine influences everyone's life, this investigation also has a wider general relevance. Questions such as can we expect a 'post physician era' with computers in place of white coated doctors, or will human staff still reassure and treat us in the clinics and wards of the future, arise. Pinch's comment (1993) that "This debate, unlike many of the rather esoteric areas of science studied in second generation SSK, has consequences for us all" (p. 369), seems particularly applicable to this research. This investigation has looked at several medical computers, and analysed the situation in practice. Particular attention has been given to how and why computers fit into our lives as social protheses, why in some instances they fit better than others, and the extent to which they deskill the workforce. My criticisms of the AI approach stem from AI workers' neglect of the processes of socialisation that people undergo in order to participate in social life - in contrast to this AI workers emphasise the formal methods of representing knowledge about the world.

2. OTHER CRITIQUES OF FORMAL APPROACHES TO PROBLEM SOLVING

Other writers have offered alternative criticisms of formal approaches - eg, Ashmore, Mulkay and Pinch (1989) discuss the Quality of Life (QOL) and Quality Adjusted Life Year (QALY) measurements devised by health economists, as an example of a formal technique for solving practical problems (see also Mulkay, Ashmore and Pinch, 1987).

They criticise this formal method by showing how the practical problem of NHS resource allocation is reformulated into an abstract economic problem by the introduction of economists' terms and definitions. The semantics of economics are a persuasive means of showing a problem to be soluble by formal methods. But, Ashmore, Mulkay and Pinch claim that the nature of the problem is altered because the inherent irrationality is removed from the problem by the reformulation (1). They explain how the subjective QOL preferences of individuals were converted into objective QALY measurements. This reformulation occurred within health economists' frameworks of how people function and experience the world. In this framework everything is reduced to a component of a QALY. As a result of the reformulation, potent solutions were produced because it was assumed that QALYs applied to large groups of patients, as opposed to individuals. Ashmore, Mulkay and Pinch stress the role of the economists' background assumptions in the process of 'QALY devising'. These assumptions are three fold - first, the assumption that the analysts' categories and the everyday experiences of the respondents correspond (2). Second, the health economists assume that peoples' preferences are stable (3). Thirdly, the health economists claim that the quantification of subjective assessments is a reasonable representation of individuals' everyday feelings (4). QALYs are presented as the preferences of ordinary people, and as beneficial to ordinary patients, but they are only a reflection of those preferences *after* filtration through the health economists' measurement techniques. These techniques reflect health economists' preconceptions and assumptions, and constitute just one way of looking at the world.

Ashmore, Mulkay and Pinch maintain that the alternative ways of looking at the social world, the social character of the research process and the social character of the knowledge claims produced should be acknowledged by health economists when they apply their expertise in areas of social policy beyond their own field of economics. The health economists should recognise that the consequences of their intervention vary according to the social context in which it is employed. Rational economic solutions work best within the confines of rational economic contexts, but when economists' formal responses are reintroduced into the irrational arena of the NHS, a complex series of exchanges ensues and the problem is moved from one context into another.

Ashmore, Mulkay and Pinch conclude that they have no interest in evaluating the epistemological status of the knowledge claims they analyse (p. 187), and that they do not intend to 'damn' the health economists, or offer their own solution to health economists' problems. Rather, their aim is to show that there are alternative ways of analysing any situation. Their analysis is essentially a moral critique of formal methods, and their reflexive stance recognises that sociologists are open to the same criticisms as health economists.

Some of the points raised by Ashmore, Mulkay and Pinch as criticisms of the health economists' formal approach have parallels with the formal AI approach highlighted in this research: AI workers must also redefine problems in their own terms; they also present their products as beneficial to ordinary patients; they often run into difficulties when their 'rational' solutions to problems (their expert systems and computerised devices) are reintroduced into the 'irrational' diagnostic arena. Many of these problems arise because the AI workers seek to replace what they see as the inherent 'irrationality' of the diagnostic process, with elegant formal solutions.

However, my stress on the socialisation argument is a different kind of criticism of the formal approach to that made by Ashmore et al. Whereas they are content to point out the multiple ways of viewing any situation rather than present any alternatives to the practical strategies of the economists, I press my case further: That there are multiple interpretations of any event from different perspectives is not at issue. However, if practical work is to commence, it is necessary to treat the world as understandable and manipulable in terms of a unitary discourse (Ashmore et al, 1989, p. 190) (5). My participatory research has allowed me to asses the situation from the insiders' perspective, and to side-step the paradoxical, multivocal nature of the world, in order to make progress. Using this strategy I will make recommendations for an improved approach to using AI in medicine and deciding where systems are likely to be most useful. Essentially these recommendations involve recognising the human contribution to the effective performance and acceptance of computerised artefacts. These practical suggestions are set out in guidelines at the end of this chapter.

The theory of behaviour-specific action and the venturi model of medical practice (adapted from Blois's work) have been used as the framework for my analysis. These theories emphasise the human role and the social factors influencing machine performance, and so fit in well to research influenced by the sociology of scientific knowledge.

In accordance with criticisms levelled at the theory of behaviour-specific action (eg. Slezak 1992), I have attempted to distinguish between the regular action and the behaviour-specific action for each of the case studies presented, so that 'predictions' about what the computers should achieve can be made. This is often difficult because of the problems of estimating actors' intentions. The next section discusses this in detail. Following this I present a summary of the main findings of the fieldwork in relation to the questions raised in chapter 4. Conclusions about the research are then drawn.

3. IS IT BEHAVIOUR-SPECIFIC ACTION OR REGULAR ACTION?

How do we decide whether a computer has stepped over the line between behaviour-specific action and regular action? Slezak (1992) says that to do this it is necessary to specify in advance what in the world of human activity is regular action and what is behaviour-specific action:

The information from the GLADYS study suggests that when doctors do dyspepsia diagnosis it can sometimes be accomplished using a series of rule based responses, ie, behaviour-specific action. At other times, unforeseeable reactions to various factors is required and regular action is involved. The same is so for the management of dyspepsia - the specific cases outlined in chapter 5 show that management is sometimes straightforward, but that regular action which responds to the unforeseeable specifics of each patient and their problems is also sometimes required. So, according to the theory of behaviour-specific action, GLADYS should be able to produce acceptable diagnoses some of the time and acceptable management decisions some of the time.

Similarly, the SOLUBILE case shows the variable nature of jaundice diagnosis. It seems that some cases can be diagnosed using a behaviour-specific action whereby doctors reproduce earlier actions, whilst others require a more unorthodox approach based on regular action. It follows that in theory SOLUBILE should also be able to produce results that doctors find acceptable in some cases, but not in all cases.

The study of the automatic blood pressure machine highlighted a task which can be completed by following rules. It seems to be a behaviour-specific action where some variability is tolerated, to which the actors are indifferent. However, people often fail to instantiate a behaviour-specific response - instead they use techniques that do not fall within the margin of tolerance of the actors. My analysis suggested that behaviour-specific action is the best way to accomplish this task and avoid variations between measurers. So, the automatic machine should produce consistently acceptable results by using a behaviour-specific method which falls within the margin of tolerance of the rule-following model.

The information from the study of the automatic ECG machine shows that ECG analysis is a two stage process. The first of these stages seems, on the face of it, to be a rule based decision about whether a black line on a white piece of paper is normal or not. On closer analysis this decision is often context specific and patient specific - made on the basis of unforeseeable details. It sometimes demands regular mental action. The second stage, which involves determining the nature of any abnormalities, is another task that may involve either regular action or behaviour-specific action in different cases. The interpretive ECG machine should be useful for some instances of screening and for some instances of abnormality interpretation.

The final case study described the HELP system. The jobs that this computer is assigned to fall into two categories. The first set of tasks are large-scale sifting of information tasks, and the isolation of specific cases which do not fit into pre-determined 'normal' categories. This task - if done by people - could be done using behaviour-specific action and HELP should produce good results using a behaviour-specific approach. The second kind of task this system does, involves responding to patient oxygen levels and in the light of this, suggesting therapy. Advising on patient therapy is traditionally seen as an area where experience and intuition are necessary qualifications for the job. It is regarded as regular action because of the wide variety of symptoms and signs that have to be considered in each different patient's case. However, as shown in chapter 9, the management of arterial oxygenation levels involves assessing only one variable. This is a very small part of the overall monitoring of the physiological body. A limited number of variations on the readings and a limited number of responses (which fall within a zone of tolerance) have been formulated in advance. Staff could deal with run-of-the-mill readings using this rule based, decision-tree type of analysis that covers all expected variations of the reading from a pre-defined normal. The 'microworld' nature of this area has prompted the designers of HELP to try to anticipate all the variations and the appropriate responses to them. However, when the patients responses fall outside those that have been anticipated and accounted for (fall outside the microworld) the doctor is called on to use the knowledge and experience gained from dealing with unexpected occurrences to rectify the problem. The machine is only designed to deal with stable sick patients, when their oxygen level variations fit a predictable model. Within this boundary, the system should work well.

On the whole these predictions about what the computers will and will not be able to reproduce have been upheld by the empirical field work: GLADYS and SOLUBILE both produce acceptable results some of the time; the BP machine consistently produces acceptable results; the ECG machine offers acceptable interpretations some of the time. The HELP system performs task type 1 acceptably, and performs type 2 tasks acceptably in some cases (6). Even though some of these machines are only performing the tasks 'acceptably' some of the time, all the systems are situated in working environments where they are used by the staff. Looking closely at how the machines and the staff interact reveals more about the capabilities of the mechanical devices:

4. HUMAN MACHINE INTERACTION

The research has highlighted the various contributions made by the people working with machines. These contributions permit the machines to blend into the social setting and make the machines look as if they do a great deal of the 'human' work. The staff deal with the real world and act as a conduit between that real world and the machine, and between the machine and the real world.

Machines do not interface directly with the world because understanding the real world beyond their area of application requires participation in the world. Socialisation and enculturation lead to the accumulation of mutual understandings. These understandings permit the competent functioning of people in the world, in a way that other people understand and respond to. We are all apprenticed to society. Computers do not have the benefit of this socialisation and apprenticeship. So they are designed in such a way that they can utilise our understandings. How does this work in practice?

In the case of GLADYS, the staff select 'suitable' patients for the computer, and the patients provide 'suitable' digitized responses for the computer to work on. The link between the wide world (which people understand) and GLADYS'S world is thus made.

In the case of SOLUBILE the doctor interviews and examines the patient, interprets the program's questions and provides specific statistical responses which the computer then works on. The doctor makes the real world of medicine into a SOLUBILE friendly set of statistics.

The automatic BP machine is provided with a fully prepared patient, with the cuff ready positioned, and the 'normal' limits agreed on and programmed. The person who does these tasks has reduced the real world understandings into a suitable shape for the machine to work with.

The automatic ECG machine is presented with a relaxed patient who has been attached to the machine in the correct manner, and whose ECG signal should be clear and sharp, thanks to the work of the ECG operator.

HELP is given a series of oxygen saturations from a stable patient. This information has been separated from the confounding mass of physiological information that the patient produces. This mass of information is dealt with separately by human practitioners.

All of these contributions that humans make to the different processes constitute 'digitization' of the world. In each case the real world is narrowed down and presented to the machine in neat discrete categories that it can work with. This narrowing down is represented in the top half of the venturi diagram. The digitization narrows down tasks so that they lie nearer to point B. They are transformed into microworld tasks where rule based approaches are suitable. (See figure 10.1 below, and chapter 4, section *1.3ii*).

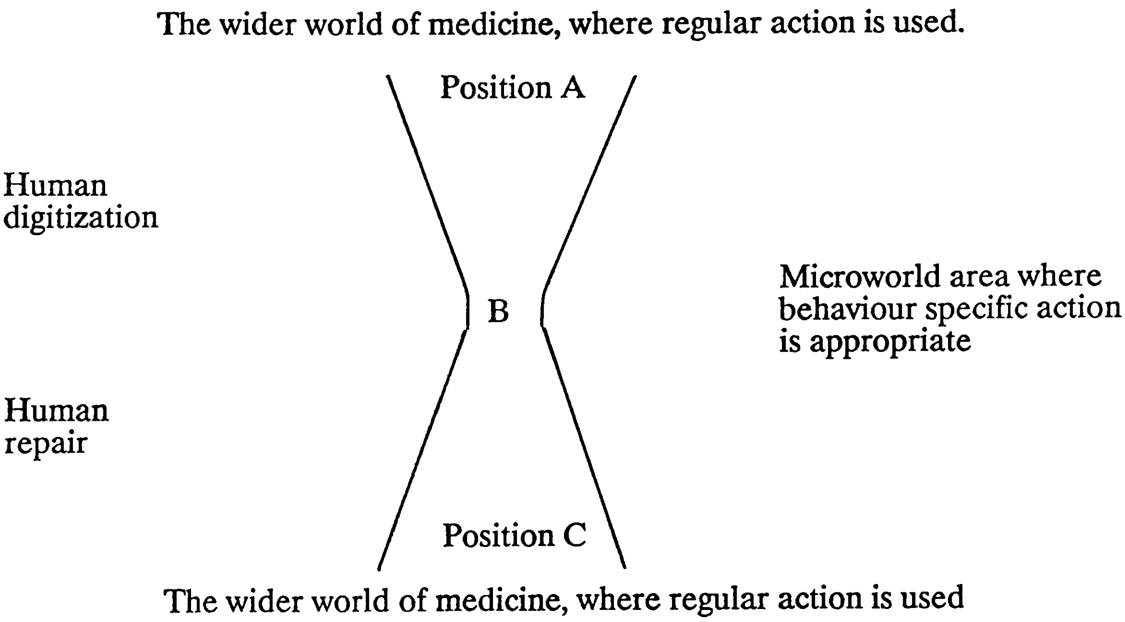


Figure 10.1 - The Venturi Architecture

This is only half of the story - people come into play at the 'output' end of the equation too. The output end occurs after the computer has performed its task in the pre-defined manner. At this stage, the human contribution makes the machines' output fit into the real world of medicine, as opposed to the microworld in which it was produced. How do people do this?

GLADYS is being used in a level 3 trial (ie, in operational form with an awareness that it is not fully proven) and its output is being compared with that of physicians. This means that physicians go about their business as usual, merely comparing their diagnoses and management strategies with those given by GLADYS. However, if GLADYS was in full operational form, a qualified practitioner would be required to stand by and oversee GLADYS. The unacceptable diagnoses and inappropriate management strategies highlighted in the specific cases in chapter 5 would have to be over-ruled and adjusted by the doctor. This is where GLADYS'S output would be 'repaired', making it fit for use in the real world.

SOLUBILE is in use in a London hospital. The doctors manipulate the list of differential diagnoses given by SOLUBILE so that they fit the case in question. The doctor decides on the relevance of SOLUBILE'S probabilities for each condition, decides whether or not to perform the tests SOLUBILE suggests, and may have to decide whether any of the diseases given a low probability need investigating. As it stands when produced, the SOLUBILE printout is only useful when a practitioner has decided what to do with it to make it useful in the real world in relation to the particular patient under examination.

Similarly with the automatic BP machine. The staff respond when the machine sounds it alarm, they also assess the numerical readings produced, in the light of the particular patient under scrutiny, and decide on the best course of action. Decisions about whether the reading is too high or too low are situated actions made by the qualified practitioner. These decisions are necessary if the machine is to fit into the world of human action.

The human repair that occurs after an interpretive ECG machine has performed its task is highlighted in chapter 8. When using the interpretive ECG machine, operators react to misinterpretations by ignoring them or switching off the interpretation facility. When the machine misses an important abnormality the technician brings it to the attention of a doctor. The role of an experienced practitioner is essential if the machine is to be used responsibly.

Finally, the ARDS protocol decisions produced by the HELP system are also subject to human repair: As shown in chapter 9 the staff decide whether or not to follow a directive. They may decline to do so if they feel the directive is inappropriate in the light of the particular patient's condition. Especially during periods when the patient is in an unstable condition the staff may decide that a directive is inappropriate.

All of these human contributions effectively repair any inappropriate computer-produced diagnoses and decisions. The charitable human transfers the machines' output from the form it took in the microworld in which it was produced, to a form that is acceptable in the real world of medicine. This part of the process is represented in the bottom half of the venturi model, between points B and C. All the machines in this study only do a small part of a bigger task - they are all dependent on the staff for assistance with the parts of the task that they cannot perform. The result of this is that the staff link the part that the machine does, to the broader world of medicine. The people working with the machines are assisting the machines, helping them fit into the environment. The digitization and repair mean that the computers become acceptable in the social environment in which they are placed, and they often appear to do much more than they really do - "most of us are sufficiently expert ... to compensate for what these artefacts cannot do." (Collins 1990, p. 215). In the case of medical 'artefacts' and machines, specialist knowledge of the medical world is required by the users of the machines in order for the inadequacies of the machines to be compensated for. The skill of medical personnel is necessary if the computers are to be employed usefully. The skill of the staff is not, then, completely subsumed into the machine.

However, the fieldwork showed that although digitization and repair can be used to fit the machine into the environment of the people, and will compensate for mechanical shortfalls, it is sometimes unnecessary. This is because the staff sometimes choose not to use the machines at all. Why does this by-passing of the machine occur?

5. WHY COMPUTERISED MACHINES ARE SOMETIMES BY-PASSED

GLADYS, SOLUBILE and the interpretive ECG machine are often ignored by the staff who are the intended users.

Although staff have agreed to participate in the GLADYS trial, they often expressed exasperation with it - this applies to the nurses who were required to help patients use it, and the doctors who were supposed to benefit from its output. The senior registrar stressed that if GLADYS was not under trial he would not be using it because it does not have access to the wider context that doctors work in when interviewing, diagnosing and managing patients. He felt that dyspepsia is too wide an area for reduction to computerisation and as a result, GLADYS gets decisions 'wrong' in the doctor's eyes, and its probability based outcomes are of little use in the real world of medicine (7).

SOLUBILE, however, is employed in a working environment. It is generally agreed that jaundice diagnosis is a 'narrower' field than dyspepsia, because a diagnosis can be made with less evidence, and the popular idea is that the computer will work well in this smaller area. But the staff complained on the one hand that SOLUBILE gave diagnoses that they did not consider to be accurate, and on the other hand about the fact that it was being used in a narrow field where they did not require any assistance. It seems that being employed in a 'narrow' field is no guarantee of success in terms of producing results that skilled staff find acceptable. The human barrier to the use of SOLUBILE is based on both these factors - the 'mechanical' problem of inaccurate suggestions, the philosophical problem of too small an area, and thirdly, the unacceptable (although interesting) method of probabilities used by SOLUBILE in differential diagnoses which mean little to practising clinicians.

The ECG machine's interpretation facility is often ignored by experienced technicians so that they do not have to assume the role of 'detector of mistakes'. They see the machine as time-consuming, often inaccurate, misleading and unnecessary. Staff at the large hospital who were well trained, do not use the machine because they recognise its technological inadequacies.

In each of these cases the acceptability of the machines to the social group is limited.

However, the automatic BP machine and the HELP system do not face a human barrier to their use. Staff are happy to use both computers. These machines produce acceptable results for a large proportion of the time. But even in these two cases, the presence of a skilled human is essential for the instances when the machine does not produce an acceptable output, or is faced with a situation it has not been programmed to respond to. It seems that three of the machines pin-pointed in this study achieve only limited success and are often rejected by users, whilst the other two are accepted - to a much greater extent - as legitimate social protheses. Explaining this difference requires close attention to the intricacies of digitization and the necessity to reduce tasks to a machine-approachable microworld form.

6. MICROWORLDS, ACTOR'S INTENTIONS AND BEHAVIOUR-SPECIFIC ACTIONS

In the cases of arterial oxygen level maintenance and BP recording the machines are more readily accepted.

First, the case of arterial oxygen level maintenance by HELP: This is a sub-task within the broader task of patient management. Through a process of 'digitization', part of the actor's world has been simplified from a complex world (overall patient management) to a microworld (oxygen level monitoring). Acting within a microworld can be accomplished by following a set of exhaustable rules, because, as Haugeland (1985) explains, a microworld is a "domain in which the possible objects, properties and events are all narrowly and explicitly defined in advance." (p. 185). Maintaining oxygen levels involves measuring and responding. As long as the measurements fall into pre-defined categories, the appropriate response can be set out in pre-determined rules and the action performed within a microworld.

Any parts of an actor's world could be reduced by digitization to a microworld format, provided that the reduction of the complex variety of the real world into invariant unambiguous categories does not represent too much of a 'loss' in terms of the actor's world and form of life. In the case of the sub-task of managing oxygen levels, what is lost by digitization is not important to the broader task of patient management. It is clear that in this field the actors' intention is usually to respond to the oxygen level measurement according to the pre-set criteria. Consequently, a machine can be programmed to reproduce this behaviour-specific action within the confines of the microworld without loss.

However, microworlds cannot be combined to add up to the complexity of the real world. This is why - in this case - when the oxygen levels fall outside the pre-specified values and rules operating in the microworld, a doctor from the real world is called on to respond using regular action. The conditions and specifications of the microworld do not apply outside the boundaries. In this broader context the intention of the actor is not so easily pre-determined. Regular action that responds to unspecifiable variations in unspecifiable ways occurs. A machine cannot be programmed to mimic this kind of response in the world outside the microworld.

Secondly, the automatic BP machine. The range of techniques that people use to record the pressure can be reduced to one particular method that relies on behaviour-specific action. 'Measuring the pressure' is reduced by digitization to a rule-following procedure - without loss to the wider complex world of 'taking the blood pressure'. The intention of the actor is to record the systolic and diastolic values in a behaviour-specific way. This task can be regarded as occurring in a microworld, where everything is determined in advance and where rules can be formulated for successful completion of the task. Behaviour-specific action will suffice in this case and the machines can be assigned to do the task.

Collins makes the point that "Those who want to substitute human labour with machines must first arrange the job so that it can be done in a machine like [behaviour-specific] way; that is where deskilling comes in." (1990, p. 221) He is suggesting that without the rearrangement of tasks into a behaviour-specific format, the role of machines is strictly limited. Both tasks described above can be arranged so that they can be done using behaviour-specific action. But the opportunities for arranging medical tasks so that they can be done in a completely behaviour-specific manner are rare. Medical treatment is, by nature, the performance of a range of tasks which span a hierarchy of levels (see Blois, 1988). As Szolovits and Pauker point out (1993) "Medical diagnosis is inately an uncertain business." (p. 171) ECG analysis, dyspepsia diagnosis, dyspepsia management and jaundice diagnosis are the tasks that the other machines in this study were designed to approach. These machines are less acceptable to users than HELP or the BP machine. So, what exactly happens when the interpretation and diagnosis parts of ECG analysis, dyspepsia diagnosis, dyspepsia management and jaundice diagnosis are reduced to microworld tasks? Close analysis shows that difficulties are encountered when these tasks are digitized in order to fit them into microworlds. To explain this it is necessary to return to the venturi model.

In Blois's original funnel description, at the time of the first doctor/patient encounter, a broad range of possibilities are open to the doctor, and a range of clinical and ordinary judgemental skills must be employed in order to approach the problem.

The narrow end of the funnel is approached when a working hypothesis has been reached, or when the range of possibilities has been narrowed down, and where a calculation or the application of highly specific expert knowledge will finish the job and produce a decision. No knowledge of the wider world beyond the microworld is needed at this point.

Monitoring oxygen levels of stable patients and responding in a pre-determined manner, and measuring the BP of patients whose cuff is already attached and who have been familiarised with the procedure, are tasks lying at the narrow end of the funnel. These tasks, which HELP and the BP machine are designed to do are near the narrow part of the funnel (or venturi). Digitizing the task to a microworld format can be achieved without loss to the actors form of life - all the variables can be predicted in advance within the boundaries of the microworld. These tasks may move up into the wider end of the funnel if unforeseen circumstances occur which require more than straightforward knowledge of the 'narrow end' microworld variety.

GLADYS and SOLUBILE have been designed to solve problems situated at the wide end of the venturi. A wide range of possibilities face the doctor that is presented with a jaundiced patient or a patient with dyspepsia. Determining the doctors' intention in advance is difficult because of the range of possibilities and the 'open' nature of the job.

Between the two extremes of the funnel architecture the automatic ECG machine is designed to do a job that probably lies somewhere in the top third of the funnel, nearer to the wide end than the narrow end. This is because the job of ECG interpretation does not occur at the point of the first doctor patient encounter - so it is not at the widest part of the funnel. But interpretation does require some of the broad skills and judgement that characterise 'wide end' tasks. It does not fall under the umbrella of narrow end tasks, where no knowledge of the wider world is needed.

Blois's point is that computers will not be acceptable when applied to wide end tasks because they do not have the wider world knowledge necessary to function at that point. Computers can be useful for narrow end tasks, which have been sufficiently sharpened by the action of humans, for a machine to finish off the job.

The designers of GLADYS and SOLUBILE face the prospect of computerising a task that lies at the wide end of the funnel. One means of overcoming this problem is to move the tasks in question from the wide end of the funnel to the narrow end. This is attempted by digitizing the information presented to the machine; ie, by reducing patient specific characteristics to a series of yes/no answers and numbers so that the individual patient is effectively 'digitised-out' of the process. But this kind of digitization changes the nature of the task: The excessive digitization that is required to move the task into the narrow part of the venturi means that essential parts of the problem and the method of solving it, are removed. The result is that neither GLADYS nor SOLUBILE are doing the same thing as a doctor would do: In the cases of both dyspepsia diagnosis and jaundice diagnosis the human doctors' task is to find out what is wrong with the particular patient presenting, in the light of symptoms, signs, history, social context, psychological factors and familial circumstances. GLADYS and SOLUBILE though, are doing something different. They are calculating the probability of various diseases being present in the light of various numerical values and yes/no answers to specific questions. They do this without taking into account the external factors that the doctors consider. The digitization is too extreme. It eliminates factors that are essential to the way the task should be approached. This inappropriate digitization I will term 'hyper-digitization'. The hyper-digitization of real world complexities into microworld categories removes all external influences. This constitutes a great loss in comparison to the real world activity of a doctor, and as a result, GLADYS and SOLUBILE often provide 'unacceptable' outputs which are rejected by the staff. However, not all diagnosis cases demand reference to these external factors - some cases are much more straightforward. Then the removal of the external factors by hyper-digitization is not crucial. In these cases GLADYS and SOLUBILE do produce acceptable results.

ECG analysis is slightly different. The machine is designed to interpret heart recordings. Both machine and human analyst are attempting the same thing - ECG interpretation. Yet the machine is often rejected because it cannot perform this task consistently. This is because the machine uses an algorithm which is a representation of the way humans interpret ECG traces. By digitizing the (assumed) human procedures, the designers have attempted to move the task down to the narrow end of the funnel. But the difference between individual patients, and the significance of each trace in the light of individual patients, cannot be anticipated and formulated into rules. The digitization of techniques into specific rules ignores such factors - they are 'digitized-out' of the mechanised process. This again is an inappropriate type of digitization, because the loss of these factors constitutes an important difference between the mechanical method of ECG interpretation and the method employed within the human form of life. The result of the hyper-digitization is that the machine often produces results that are unacceptable within the human form of life. Sometimes though, the machine produces acceptable interpretations in cases where the digitized rules are close enough to the procedures used by people, to make no difference. Again, in some cases, the losses incurred during digitization are not critical.

What arises from this discussion is the idea that reducing a human task at the wide end of the funnel to a form that a machine can perform involves moving the task to the narrow end of the funnel. Attempts to achieve this are made by hyper-digitization of the task and exclusion of many external factors. As a result, the machine often ends up doing a different job to the human and this is why computers and people sometimes produce divergent results when assigned to the same job. The hyper-digitization removes the context based factors and fits all variables into pre-defined categories. In the cases of GLADYS and SOLUBILE, hyper-digitization changes the nature of the task that is left for the computer compared to the task that the doctor would have done. In the case of ECG analysis, it results in the machine working with a set of interpretation algorithms which have digitized-out some of the variables that human staff assess when interpreting ECGs.

However, when we consider the jobs that HELP and the automatic BP machine do, the situation is different. These tasks are positioned at the narrowest part of the venturi - they can be performed with little reference to patient specific or contextual influences. The digitization required to make these tasks computer-suitable is minimal. Whether a person or HELP performs the task makes no difference - it remains the same task - measure and respond to oxygen levels. Similarly with the BP machine, the task is detect and record blood pressure pulses, which is what a human recorder also does. They are microworld tasks. Ignoring the external factors - as the machines do - does not constitute a great loss in comparison with the world of the human actor. The context and patient related factors are taken into consideration **after** the machine has performed its analyses. Then the staff decide on the validity and relevance of the machine's output in relation to the particular patient under investigation.

To summarise these ideas: The computerised machines are useful when assigned to tasks which lie in the narrow section of the venturi, where patient specific characteristics are less important and can be 'digitized out' of the process without significant loss to the task at hand. But some tasks belong at the wide end of the funnel and are greatly influenced by external factors at many stages during the procedure. Hyper-digitizing these kinds of tasks to fit into a microworld format, and artificially 'moving' them into the narrow end of the funnel means that these characteristics are ignored. This constitutes a great loss to the actors world and the characteristics of the task. As a result these machines are working with a different set of information, sometimes even performing a different task, and so sometimes produce different results to the staff. This explanation goes some way towards showing why some computerised systems produce acceptable results and others do not. It is to do with where they lie in the venturi, and the effects of moving them into the narrowest section of the venturi in order to reduce them to microworld tasks where behaviour-specific action will reproduce the actor's intention.

7. CONVERTING MICROWORLD OUTPUTS INTO WIDER WORLD TERMS

The human contribution to each of these processes that makes the mechanised output acceptable in the real world, is represented in the bottom section of the venturi. Here, staff repair the output created at point B and move it to point C. Blois (1980) recognises that a great deal of work is often done before the task is given to the computer, but he does not refer to this work that humans do **after** the computer has performed its analysis.

In the cases of GLADYS, SOLUBILE and the ECG machines, the extent of the repair required can be extensive - particularly in instances where the hyper-digitization was inappropriate. In some cases, the digitization and repair required are too time consuming and the rewards in terms of assistance to experienced staff, are too small. These staff may then reject the technology, recognising the technical shortfalls.

However, I have also shown that when inexperienced staff are presented with the same technology that experienced staff reject, it is possible that they will not recognise the inherent inadequacies of the technology. For example, less experienced staff are far more inclined to use the ECG machine interpretation facility, because they do not recognise the inadequacies of the device. This is because less experienced staff work in a slightly different form of life to highly skilled staff. For the inexperienced, the hyper-digitization of the task constitutes less of a 'loss' because they are at a different starting point to the experienced personnel. They do not themselves use a full range of regular actions to respond to each new case, but seem to function at a less developed level. The technicians may be functioning at a rule-based behaviour-specific level. They lack the tacit knowledge and expertise required to make unorthodox decisions based on individual factors in each new case. Their methods lie near the narrow part of the venturi. So they do not see a computer which functions at the narrow section level, as problematic. A clear example of this is detailed in chapter 8, where the untrained ECG technicians and doctors unfamiliar with ECG evaluation were content to accept the interpretation offered by the computerised machine. They did not recognise the inadequacies of the machine nor the necessity to repair its output, because its method of operation was in fact similar - or perhaps even better than - their own method.

8. AN ALTERNATIVE INTERPRETATION

This explanation of where machines work best and how likely they are to prove acceptable, emphasizes the interaction between the machines and the people using them, and stresses the importance of human digitization and repair. My account shows that the efficacy of the machines - in the eyes of the users - depends on different levels of intervention by those users.

An alternative view is that the efficacy of the machines and their level of acceptability is (or will be) determined by processes of negotiation between people. The social constructivist perspective (outlined in chapter 1) presents this negotiation as the crucial factor in the determination of the adequacy and acceptability of technology.

In the case of the machines highlighted in my case studies, negotiations of this sort about the 'adequacy' of medical machines involve users, designers and purchasers - ie, medical staff, AI workers and administrators - each party using negotiating 'tools'. Mackenzie (1989) suggests that the testing phase is a vital stage in deciding where and how well machines work, and what their characteristics are (see p. 12). It is during the testing and trials of medical machines in working environments that users make decisions about machines, whether they 'work', and whether they are 'adequate' and suitable for everyday use. These decisions to accept or reject are influenced by social processes of negotiation. The 'usefulness', or perceived usefulness, of the machines emerges from these negotiations.

Anderson's (1992) interpretation (see chapters 1 and 4) fits in well with the social constructivists' analysis of the processes by which machines are determined to be adequate or not. Anderson explained the rejection of diagnostic technology in a Melbourne hospital in terms of a power struggle and debate between opposed groups of diagnosticians and scientists - the diagnosticians eventually ensuring the rejection of the technology. Anderson accepted the results of an ill-conceived clinical trial which he believed demonstrated the usefulness and accuracy of the devices. His analysis is weakened because he disregarded the perceived usefulness and adequecy of the machinery to the users, which arose from the debates.

The philosophical issues associated with AI are illuminated by the practical applications of the technology, and the practical experiences of users cannot be ignored. Analyses such as that provided by Anderson do not seem to recognise the close proximity of the philosophical issues and the practical issues.

It is clear that the decision whether to accept a machine - on a trial basis, or more permanently - *is* influenced by social factors, argument and consensus (8), and the usefulness of social constructivist perspectives is not at issue. However, my contribution - which highlights salient points concerning human-machine interaction at the point of use and the users' perception of the usefulness of the devices - is an important component of constructivist interpretations of events (9). The guidelines provided at the end of this chapter show how closely related the perceived adequacy of the devices (in the eyes of users) and the acceptance or rejection of the devices are.

9. DISCUSSION

The picture presented here is that some machines are extremely useful for doing limited, microworld tasks where behaviour-specific action produces acceptable results.

Within strictly defined boundaries these machines can replace people. However, it has also emerged that when machines are applied to tasks that do not readily fit into the microworld format problems can arise.

When jobs demanding regular action and reference to external factors are reduced to a microworld format, the external factors are ignored. 'Creating' microworlds in this fashion is not always successful because in so many cases the external factors are crucial to the performance of the task.

Inappropriate or hyper-digitization of tasks by designers (as in the ECG machine), or by users (as in the cases of GLADYS and SOLUBILE) does enable the machines to approach the task. But it may result in the machines approaching a different task to the staff, or taking note of different factors to experienced people. This can lead to unacceptable output and rejection by the users.

Microworld tasks from the middle of the venturi where behaviour-specific action is sufficient are rare in medicine. Transferring wider world tasks into a microworld format is fraught with difficulties. Medical expert systems and other computerised artefacts can be equipped with basic knowledge about sub-sections within medicine. All independent pieces of information that stand alone, and apply to all patients in all situations, can be programmed into a machine. This is the extent of the system's 'knowledge'. This is the kind of knowledge needed to solve problems lying in the middle of the venturi. At the wider end of the venturi influences from other spheres of medicine and from the real world play a part in the decisions made about individual patients in different situations. Then doctors draw on their tacit knowledge, gained from being in the wider world, in order to approach problems. Computers do not have access to this level of tacit knowledge because they are not participants in our form of life. There is then, a gap between the two levels of knowledge. This gap can be narrowed if the machine is equipped with more and more complex rules of behaviour. These can be formulated in retrospect after any event. But rules for coping with next week's unique patient cannot be formulated in advance. Ramifying rules allows a computer to come a little closer to the mimicking of human regular action. But rules do not encompass the never ending variety of responses that humans display when they are doing regular action. Accessing the full range of human knowledge cannot be achieved without participation in the human form of life, and interaction in the human collectivity. Machines do not participate or interact. Socially gained collective knowledge of the type described by sociologists of scientific knowledge is necessary for performing in the real human world. Machines do not have access to this knowledge. So, machines do not perform at the same level as socialised humans.

This discussion is based on empirical results. Have the designers and manufacturers of the 1990s altered the goals of their research in view of their acknowledgement that "the dissemination and use of AIM systems has remained minuscule."? (Szolovits and Pauker, 1993, p. 178). Since the early pioneering days, Szolovits, often publishing with Schwartz, Patil, and Pauker, has been a prominent researcher in the field of AI in medicine. Schwartz has now abandoned the field, perhaps losing interest or recognising the substantial hurdles to AI in medicine (10), but Szolovits and Pauker are still active.

10. THE VIEW OF AN AI RESEARCHER

In late 1991, Professor Szolovits and I discussed developments in the field of AI in medicine, and the changes that have occurred since the early 1970s. He explained that research had originally been prompted by three goals. The third of these was

> "to build the practical systems that we built here, and in some senses I think that's the one at which it has been least successful."

This lack of success has dampened initial enthusiasm:

> "by comparison with the enthusiasm that I and many of my colleagues felt fifteen years ago we certainly have not achieved as much as we thought we would at the time."

The reason for the lack of success early on was, in his opinion, the naive model of how computers would be used in medicine which the designers worked toward: They saw a 'doctor in a box', consultant type of system as the best application of the burgeoning technology. This model worked on the underlying principle that if a patient exhibits symptoms and it looks as though they have a disease, then they do have that disease. But this kind of reasoning ignores overlapping symptoms, multiple diseases, opposing symptoms, evolving illnesses and the changes that occur as physicians intervene. Szolovits felt that dealing with these complexities of diseases would require

> "thinking like a physiologist, and considering the mechanisms whereby the body generates these symptoms."

Computerising this 'pathophysiologic' reasoning was seen as the way forward - an improvement on the method of merely considering associations between diseases and symptoms. Szolovits explained that this model was much more useful, but was difficult to operationalise in practice - much of medicine is not understood to that degree of detail, the programs could become enormously complex and

> "a lot of stuff happens in medicine for which we just don't have those models."

In view of this, emphasis had then shifted to

> "relatively self-contained domains in which you can do the job. My guess is that there are many such [domains] in medicine and I think a lot of such programs will come to the fore over time ... Remember, none of these [domains] is totally limited ... and the rules in the program may be inappropriate. But in the average case if there is some human pre-filtering to make sure that you don't buy it in inappropriate circumstances then it is probably OK."

He used ECG interpretation machines as an example of a machine working in a domain where 'human pre-filtering' allowed the machine to operate. He also recognised that human filtering was necessary after any mechanical interpretation, saying

> "If its diagnosis is that a man is pregnant and that accounts for his problem, the user should say 'Gee, I wonder what has screwed up in this program.'"

Szolovits described another 'limited domain' program he had been involved with, designed to advise on treatment for diabetic ketoacidosis - an acute and relatively isolated condition. This machine has undergone trials at a Boston hospital where it apparently performed as well as house physicians. But the staff stopped using it after the trial. His explanation for this was

> "I think if you look at a big city hospital like this and ask them how often do they think that they badly blow a diabetic ketoacidosis case, they will tell you 'not very often'. It is very difficult to get people motivated for something that works in a narrow domain."

So could this kind of system be more usefully employed in environments staffed by less qualified personnel? He immediately suggested two problems with this. Firstly, the staff must recognise the basic nature of the main problem in order to decide to use the program. Secondly, many programs rely on the doctor inputting detailed information from tests and examinations. Both these points emphasise the need for a minimum level of skill and knowledge, without which the program is unusable.

In conclusion he suggested that machines and people often use different strategies and in some instances the program will do things better than a person. He used the example of

> "trying to integrate some sort of complex probabilistic model where you have a lot of uncertainty based on tests that give uncertain results and you are trying to figure out if the patient has some heart disease or something. It is clear that programs are much, much better at that than people."

He felt that the optimum use of machines is in conjunction with staff where they serve as error detectors, and offer advice about what you shouldn't be doing rather than what you should be doing. These machines would effectively say to the doctor

> "I don't know if this is really an error but it looks serious enough to me as a computer program that I thought I would bring it to your attention. **You** decide what to do about it."

Szolovits admitted that little practical success had been achieved in terms of producing large scale diagnostic programs. He was aware that

> "If we were judged by a sort of corporate criteria of have we built the product and has it gone to market and has it made money, we would be in very deep trouble. [But] our score card is in terms of scientific progress and not in terms of deliverables ... our reviewers are other scientists who are used to working in the laboratory on fundamental questions that take a long time to answer."

To summarise, this researcher acknowledged the fundamental problems of designing and building large scale systems. The impracticalities of using pathophysiological reasoning that he talked about and the notion that

> "If you learn medicine that doesn't make you a good doctor. To be a good doctor you have to practice at it,"

show that he is aware of the philosophical difficulties involved in codifying the practices and the tacit skill used in everyday medical practice. He felt that systems working in smaller domains will be easier to build and deploy. But he recognises that it is in the more limited domains that doctors need least help (see above discussion of the ketoacidosis program). Szolovits advocated the use of systems as error detectors, with the human input and repair an essential part of a process, whereby the system ultimately says to the doctor "You decide what to do about it." This represents a substantial shift to the position he took twenty years ago.

However, in a recently published paper, Szolovits and Pauker (1993) present a slightly different picture. They acknowledge that the goals of expert system designers have evolved: Initially the designers had aimed at programs which "focussed on making a correct diagnosis and demonstrating expertise similar to that of experienced clinicians", (p. 177) in order to redistribute medical expertise. Szolovits and Pauker suggest that the change in aims is based on a change in medical consumers' demands. Instead of demanding expert quality of care for all, consumers are now calling for uniform accessible care at affordable prices. Consequently, "The expert system developer may no longer be able just to simulate the behaviour of experienced clinicians" (p. 177), but must respond to these new consumer demands. This is the reason why "Narrower more specialised systems may have a more compelling, if limited impact." (p. 176) The authors recognise that limited domains are problematic in so far as they do not offer a solution to the broad needs of clinical medicine. In order to solve this dilemma, work is needed that will create common frameworks, so that specialized systems can work together to produce a usable useful whole (p. 176). The implication here is that tackling the broader problem of computerising bigger chunks of medicine can be achieved by creating lots of specialised 'narrow' systems and linking them together.

In 1991 the philosophical and technological difficulties of the large scale approach were recognised as important. Now, Szolovits and Pauker are presenting a change in consumer needs as the reason that the large scale approaches were abandoned. In this paper, the inherent problems of creating artificial experts for use in the medical field are presented as soluble by the use of a different technological fix - the combination of specialized systems.

11. CONCLUSION

Rather than accepting the 'performance figures' of computers as a starting point, and explaining why they are still sometimes rejected, I have approached the question from a different angle. The insider perspective presented here has highlighted the problems of using computers (which apparently work) in practical settings. This methodology has revealed features about the usefulness of computers, their performance in practice, the role of staff and the extent of deskilling, which other analyses (eg. Anderson, 1992) have ignored.

In order to be useful and usable, medical machines must be confined to microworlds where success can be accomplished using a rule based technique, or behaviour-specific action. But some areas of medicine cannot be reduced to microworld tasks. Much of medicine relies on tacit knowledge and the application of regular action in unanticipated circumstances. So, machines often produce results that do not fit into the world of regular action. Human 'repair' is required to fit the microworld outputs into the real world setting. The role of an experienced human remains essential when computers are used.

In some environments where staff are less than expert, the repair required to fit machines' output into the real world of medicine does not take place. The staff are happy to work with the machines. Evidently, the recognition of mechanical inadequacies that need repairing depends on the specific form of life of the workers in each department. In some cases, inappropriate hyper-digitization of a regular action task into a microworld format does not constitute a loss to the world of the human actor, because not all human actors operate in exactly the same form of life. They do not all have the same level of skill, or ability to recognise the inadequate mechanical performances that may result, because they do not perform the task in the 'optimum' manner themselves. Some inexpert staff see no need to repair the results given by the machine. The acceptability of any device depends on the cultural context and form of life of the users. What is 'acceptable' in one context is unacceptable in others.

Frequently the result is that some users reject the technology, others accept it because they themselves lack expertise in the field. If machines are to play a useful role, provide acceptable information to experienced staff, and provide safe information to less experienced staff, two options are open:

Either (1) we change the areas in which the machines are used, and confine their operation to tasks already positioned at the narrow section of the venturi where behaviour-specific action suffices, or,

(2) we change the way we use the machines, and insist that 'expert' computers are used only by expert staff.

Both these suggestions depart from the images of a post-physician era: The first option means that machines will not embrace any of the reasoning and diagnostic tasks undertaken by medical staff working in normal practice. The second option means that the role of the qualified doctor, nurse or paramedic remains central. Neither course of action permits the replacement of expert staff by machines. Neither course of action allows machines to appropriate the entire spectrum of skilled human action. So, the implications for staff are not as extreme as they may have feared. Their role remains essential. The inarticulable nature of human knowledge, the unspecifiable nature of regular human action and the positioning of many tasks at the wide end of the venturi, ensure that the division between human actors and non-human artefacts remains intact and essential.

At this point it is useful to consider Collins's prediction about the conditions under which we may come to think of machines as 'thinking' and fulfilling Turing's prophecy (Collins, 1990, p. 222). Collins says that there are four ways in which we might move toward such a state of affairs: Either (i) machines get better at mimicking us; (ii) we become more charitable to machines; (iii) we start to behave more like machines ourselves (iv) our image of ourselves becomes more like our image of machines.

Option i seems unlikely, given the inarticulable nature of much of medical knowledge. The regular action used cannot be pre-determined and programmed, and machines cannot achieve human action by reproducing more behaviours.

Option ii is increasing, as we repair machines' output with little thought. However, in the case of medical machines I have shown that demanding more charity and repair from users can lead to the by-passing of the machines, or their inappropriate use by inexperienced staff. It does not lead to unconditional acceptance of machines by all users.

Option iv, whereby our image of ourselves becomes one of an inefficient machine, seems unlikely in this field. Differences between the way humans act and the way machines mimic those acts are evident, and usually recognised by users. However, this image of humans as inefficient machines is more likely amongst inexperienced staff who view the machine as an expert. This must be guarded against.

Option iii is the most likely means of increasing the level of computerisation of medical tasks. Only by reducing real world regular action to behaviour-specific microworld action, can computerisation progress.

This kind of computerisation can be useful to staff, as the BP machine, and the HELP system show. But combining microworlds in order to approach the bigger problems of computerisation is not feasible. Some parts of the real world outside the microworlds cannot be successfully reduced to a microworld format. GLADYS, SOLUBILE and the ECG machine show the difficulties that are involved. Szolovits and Pauker's notion of creating frameworks of limited systems as a means of computerising bigger areas of medicine, is flawed.

The use of computers in medical environments does not challenge the SSK conception of knowledge as social. Neither do these computers challenge the essential role of tacit skill in the performance of medical tasks. The computers rely on human knowledge that has been gained in the social world, and human tacit skill, to fit them into the wider world of medicine. The divide between humans and non-humans (machines) remains clear. The challenge made by radical reflexive sociologists, to this demarcation is misplaced.

Computers in medicine achieve optimum usefulness in microworlds where behaviour-specific action produces the desired outcome. Tasks which can only be reduced to microworlds by unacceptable levels of hyper-digitization are unsuitable for computerisation. Designers need to concentrate on isolating 'naturally' occurring microworld tasks, if they intend to advance the level of computerisation. They need to identify tasks - or parts of tasks - that already lie in area B of the venturi. In these areas staff rejection is rare, inappropriate use is rare and computers provide most benefit to users. Technically, such systems are reliable and useful because the underlying philosophical problems which plague much of AI are avoided

12. IMPLICATIONS FOR THE INTRODUCTION OF EXPERT SYSTEMS INTO MEDICAL SETTINGS

This theoretical analysis explains why some medical machines are more acceptable than others. This analysis can be developed into a set of practical guidelines for deciding where medical systems will be most useful and where they may fail to meet expectations:

My research has shown that computers should ideally be applied to microworld tasks where behaviour-specific action is sufficient to achieve the intention of the person who previously performed the task. This is where the machines are most beneficial in the context of real medicine. But such microworld tasks are rare in medicine. Exceptions are the tasks performed by the automatic BP machine and the oxygen monitoring facility of HELP (11). These are narrow-section-of-the-venturi tasks. However, machines are also being designed for use in areas that do not fit the microworld model, ie. wide-end-of-the-venturi tasks. In these cases I have shown that hyperdigitization is required to reduce the medical world into microworld format, and substantial repair is required in order to fit the microworld output of the machines into the real world of medicine (12). The 'adequacy' of the machines largely depends on the intervention and ability of the human user.

Two questions for a potential user of a computerised machine arise from this - firstly, what kind of task is the machine designed to perform? Is it a narrow part of the venturi task - a ready-made microworld where all the variables that the machine may encounter can be predicted in advance and responses to those variables pre-set? Or is it a wide end task which has to be hyperdigitized by fitting wide world variables into microworld categories?

Secondly, what kind of staff will be using the device? Will they be highly skilled experienced staff, familiar with the nuances of unusual cases and more mundane cases? Or will they be junior staff and paramedics with minimal experience? Will they be able to recognise when hyperdigitization has occurred, when it is acceptable or when it has changed the nature of the task, removed some relevant factors and created a situation whereby the machine is approaching a different task to the people? Will the users recognise an output that needs to be repaired? How well the machine will be seen to work - how 'useful' it is perceived too be - depends largely on the skill and professional form of life of the users. At one end of the spectrum a machine may be seen as indispensable by novices, whilst other highly professional staff may view the same machine very differently. A potential purchaser or user of a computerised diagnostic machine has to decide whether the instances of 'unacceptable' performance are too frequent for the particular intended users to deal with (13).

A further consideration is the role that the users will play in the operation of the machine. Here the two issues of type of task and type of user must be considered in tandem:

First, consider **wide end of the venturi tasks:** If devices operating at this level are used by inexperienced staff they may not recognise that the output may be distorted because of the hyperdigitization, and that it often needs to be repaired in the light of the particular patient under scrutiny or the particular circumstances of the case. In these cases, inappropriate decisions made by the machine may be taken at face value. However, if these devices are used by experienced personnel they may by-pass the machine altogether because the hyperdigitization is too time consuming (14), the machine gives inaccurate, misleading results (15), the machine gives percentage possibilities that are considered to be clinically interesting but not useful (16), or because the machine operates in too narrow an area where they do not feel they require assistance (17).

On the other hand, consider **narrow end of the venturi tasks:** If machines operating at this level are used by inexperienced staff, they need to learn how to recognise cases that fall outside the machine's microworld, and how to respond to them. They also need to learn how to perform all the digitization and repair required to embed the machine into the professional situation into which it is introduced (18). Similarly with experienced staff - they must act as overseers of the machines, and still perform many of the tasks they did prior to the introduction of the machine. Again in these cases, experienced staff may choose to by pass the machine and do the job themselves, unless the machine produces more consistent results than a group of people (19).

In all four scenarios above, a certain degree of expertise on the part of the staff is required. The human contribution to the functioning of the devices is unavoidable and essential.

The type of task the machine is to be applied to, the social/professional context in which the machine is to be deployed, the skill of the staff and the availability of alternatives to the machine are all important contributors to the social negotiations of adequacy and acceptability. Each of these issues influences decisions about whether a machine will be useful, and whether or not to deploy a computerised device. All the machines in my study were felt to work well some of the time; some of them were believed to work well most of the time. Whether the machine is adequate in any specific case depends on the variety of factors outlined above in relation to the particular case under scrutiny, and so predicting where machines will be useful and where they will fail to meet expectations is fraught with difficulties. However, the guidelines in table 2, (overleaf) may be useful to potential users attempting to estimate the potential usefulness and acceptability of a machine:

TABLE 2 -

Guidelines for potential users of medical expert systems

1. Does the task the machine is designed for exist in a microworld at the narrow section of the venturi diagram?

2. Or is the task situated at the wide end of the venturi where wider world influences apply?

3. How much digitization are the staff required to perform before the machine can approach its task?

4. How often will the staff be required to repair the output of the machine?

5. What environment and professional form of life is the machine targeted at?

6. How skilled do users need to be to operate the machines?

7. Will new training regimes be required in order for staff to offer the machine suitable input data, and in order for the staff to be able to repair outputs?

8. What benefits does the machine offer to the staff, and do these benefits outweigh disadvantages?

In essence, do the machines offer 'logical benefits' to their users? Evans (1991) explains that a logical benefit occurs when a process that is not efficient when performed by people becomes efficient through the use of a computer, and improvements in the process of care ensue (see chapter 9). All medical machines should offer the user this kind of benefit (20) as there is little to be gained from computerising a task if the result is less efficiency. So, the major consideration of potential purchasers should be whether the machine will be perceived as offering logical benefits to each set of users in their particular context.

The optimum use scenario is for the machine to operate in a microworld, using behaviour-specific action, and improving the efficiency of a process or procedure. Deviations from this optimum need to be assessed in the light of prevailing circumstances, and in view of the available alternative. Clearly what is an 'acceptable' situation for the deployment of a machine is different in a clinic staffed exclusively by auxiliary nurses, than in a teaching hospital staffed by experienced nurses, doctors and paramedics.

A vital point is that however 'useful' a device proves to be, its output is still *applied* in practice by an experienced professional undertaking situated action in the real world of medical practice. The interaction of people and machines is the key to the 'adequate performance' of mechanical artefacts.

FOOTNOTES TO CHAPTER 10

1. They say that "although the economists are, in one sense, adopting the managers' problem and producing a clear rational solution, they are also, at the same time, redefining and altering the nature of the problem itself." (Mulkay et al, 1987, p. 545)

2. Ashmore, Mulkay and Pinch suggest that in practice the research process and the act of measurement alters the parameters being measured, and that categories and real experiences do not then correspond.

3. Whereas Ashmore, Mulkay and Pinch suggest that respondents react to the socially defined data-gathering situation and their 'preferences' alter. So, the relationship of what is elicited to what occurs in everyday life is not straightforward. This is similar to the notion that opinion polls reflect voters preferences. Critics suggest that what is elicited 'on the spot' from voters is different to what occurs in everyday life when the real vote is made.

4. However, Ashmore, Mulkay and Pinch argue that QALYs do not reflect pre-existing values, but arise from the peculiar interaction of the economists and lay people, and the economists' interpretation of that interaction, and their production of the QALYs.

5. Collins suggests that as sociologists we need to think as social realists - temporarily abandoning the relativist perspective and treating the social world as real and something about which we can have sound data (1981a, p. 216-217). He favours dismissal of the self critical line, as it stands in the way of progress.

6. I am using the term 'acceptable' to refer to those interpretations, diagnoses and suggestions that the machines produce and which staff are in agreement with

7. Furthermore, this doctor felt that the prime objective of GLADYS, to reduce the number of referrals made by GPs, was misplaced.

8. In the course of the fieldwork it became clear that many of the decisions to test or use computerised machinery were made after 'negotiation' between designers and users. For example, the gastroenterology consultant became involved in the GLADYS trial after 'persuasion' by his friend and colleague on the Glasgow team. The Glasgow workers also provided a free word processor to the Cardiff team, and since the Cardiff consultant had delegated all GLADYS type tasks to his senior registrar, this proved to be persuasion enough. Similarly with interpretive ECG machines - it is recognised that manufacturers deals which offer the interpretive models at the non-interpretive price, combined with the glossy hi-tech output of the newer models, are reason enough to buy the new model. Another social factor is the 'slush' money, which is often provided by public fundraising on condition that the most up-to-date, computerised model of machine is purchased. Clearly the social factors do, to a certain extent, influence the decisions made.

9. Whilst acknowledging the influence and usefulness of this work in SSK, a thorough analysis of the social construction of the adequacy of medical

devices was not the objective of this study. Such an analysis would require a different methodology and a different thesis would have emerged.

10. Reproduced overleaf is the letter I received from William Schwartz when I suggested he met with me to discuss the developments in this field.

11. The intention of people doing these jobs is to measure the pressure of the blood in the brachial artery and to measure and suggest responses to a finate number of oxygen values. This is what the machines are also designed to do.

12. GLADYS, SOLUBILE and the ECG machines fit into this category of wide-end-of-the-venturi tasks, which have been moved into the narrow end by hyperdigitization.

13. For example, the interpretative ECG machine was deemed acceptable by most staff at the small Welsh hospital. Whereas the same machine was regarded with suspicion at the University teaching hospital. The differences arose from the different levels of skill of the users, and because the alternative at the smaller hospital was very often no expert ECG interpretation.

14. As is the case with GLADYS and SOLUBILE.

15. As the experienced ECG technicians found with the interpretative ECG machine.

16. As in GLADYS and SOLUBILE.

17. As was suggested by the users of SOLUBILE.

18. As the proxy strangers using the automatic BP machine illustrated, this 'embedding' of the machine is by no means straightforward.

19. As the BP machine and the HELP protocols do.

20. A computer which searches large patient databases - eg. the infectious diseases and drug alerting programs of the HELP system, detailed in chapter 9, is a good example of a system providing 'logical benefits'.

TUFTS UNIVERSITY

WILLIAM B. SCHWARTZ, M.D.
Vannevar Bush University Professor
Professor of Medicine

School of Medicine
136 Harrison Avenue
Boston, Massachusetts 02111
(617) 956-6702

October 29, 1991

Joanne Hartland
Science Studies Centre
University of Bath
Claverton Down, Bath BA2 7AY
United Kingdom

Dear Ms Hartland,

I am afraid that a visit with me would not be of much use because I left the field of Artificial Intelligence several years ago. I am not any longer in touch with the developments that would be of interest to you but am sure that Professor Szolovits will be of great assistance.

It's nice to know that our work in the field has been a stimulus to your thinking.

Sincerely,

William B. Schwartz

William B. Schwartz

BIBLIOGRAPHY

Alton E W F W et al (1990) "SOLUBILE' : Decision making in the diagnosis of Jaundice' Int J Biomed Comput 27:47-57

Anderson W (1992) 'The Reasoning of the Strongest: The Polemics of Skill and Science in Medical Diagnosis' Social Studies of Science vol 22:653-684

Ashmore M, Mulkay M and Pinch T (1989) 'Health and Efficiency: A Sociology of Health Economics' Open University Press, Milton Keynes

Atkinson P A (1977) 'Mock-ups and Cock-ups: The Stage-Management of Guided Discovery Instruction' in Woods P and Hammersley M (eds) 'School Experience' Croom Helm, London

Atkinson P A (1988) 'Discourse, Descriptions and Diagnoses: Reproducing Normal Medicine' in Lock M and Gordon D R (eds) 'Biomedicine Examined' Kluwar Academic Publishers, Dordrecht, The Netherlands

Banta et al (Sept. 1985) 'New Cardiograph Family with ECG Analysis Capability' Hewlett Packard Journal:23-28

Barnes B (1974) 'Scientific Knowledge and Sociological Theory' Routledge and Kegan Paul, London

Barnes B (1983) 'On the Conventional Character of Knowledge and Cognition' in Knorr-Cetina K D and Mulkay M (eds) 'Science Observed' Sage, London

Barnett G O (1982) 'The Computer and Clinical Judgement' New Engl J Med 307 (8):493-494

Berger P L (1967) 'Invitation to Sociology' Anchor Books, NY

Blois M S (1980) 'Clinical Judgement and Computers' N Engl J Med 303 (4):192-197

Blois M S (1983) 'Conceptual Issues in Computer-Aided Diagnosis and The Hierarchical Nature of Medical Knowledge' Journal of Medicine and Philosophy Vol 8:29-50

Blois M S (1985) 'Evaluating RECONSIDER, a computer program for diagnostic prompting' Journal of Medical Systems 9 (5/6):379-388

Blois M S (1986) 'Diagnosis versus Diagnostic Programs' 'Medinfo 86' :225-227, Elsevier Science Publishers, North Holland

Blois M S (1988) 'Medicine and the Nature of Vertical Reasoning' N Engl J Med, 318 (13):847-851

Blois M S et al (1981) 'RECONSIDER: A program for generating differential diagnoses' Proc 5th Annual Symposium on computer applications in medical care.

Bloor D (1973) 'Wittgenstein and Manheim on the Sociology of Mathematics' Stud. Hist. Phil Sci, 4 (2):173-191

Bloor D (1976) 'Knowledge and Social Imagery' Routledge and Kegan Paul, London

Bloor D (1981) 'The Strengths of the Strong Program' Phil of the Social Sciences 11:199-213

Brante T and Hallberg M (1991) 'Brain or Heart? The Controversy over the Concept of Death' Social Studies of Science vol 21:389-413

Braverman, H (1974) 'Labor and Monopoly Capital' New York, Monthly Review Press.

Buchanan B G (Jan. 1986) 'Expert Systems: Working Systems and the Research Literature' Expert Systems 3(1):32-51

Burke J P et al (1991) 'The HELP system and its application to infection control' Journal of Hospital Infection vol 18, Supplement A:424-431

Callon M (1986) 'Some Elements of a Sociology of Translation: Domestication of The Scallops and The Fisherman of St. Brieuc Bay' in Law J (ed) 'Power, Action and Belief: A new Sociology of Knowledge?' Routledge and Kegan Paul, London

Callon M and Latour B (1992) 'Don't Throw the Baby Out with The Bath School' in Pickering A (ed) 'Science as Practice and Culture' University of Chicago Press, Chicago

Card W I (1967) 'Towards a calculus of medicine' in Medical Annual John Wright and Son, Bristol

Clark A (1989) 'Microcognition: Philosophy, Cognitive Science and Parallel Distributed Processing' MIT Press, Cambridge Mass

Classen D C et al (1991) 'Computerized Surveillance of Adverse Drug Events in Hospital Patients' JAMA vol 26 (20):2847-2851

Collins H M (1975) 'The Seven Sexes: A Study in the Sociology of a Phenomenon, or the replication of experiments in Physics' Sociology, 9:205-224

Collins H M (1981a) 'What is TRASP? The Radical Program as a Methodological Imperative' Phil of the Social Sciences, 11:215-224

Collins H M (1981b) 'Stages in the Empirical Programme of Relativism' Social Studies of Science, 11:3-10

Collins H M (1983) 'An Empirical Relativist Programme in the Sociology of Scientific Knowledge' in Knorr-Cetina K D and Mulkay M (eds) 'Science Observed' Sage, London

Collins H M (1984) 'Researching Spoonbending: Concepts and Practice of Participatory Research' in Bell C and Roberts H (eds) 'Social Researching: Politics, Problems, Practice' Routledge and Kegan Paul, London

Collins H M (1985) 'Changing Order' Sage, London

Collins H M (1987) 'Expert Systems, Artificial Intelligence and The Behavioural Coordinates of Skill' in Bloomfield P (ed) 'The Question of Artificial Intelligence' Croom Helm, Kent

Collins H M (1990) 'Artificial Experts - Social Knowledge and Intelligent Machines' MIT Press, Cambridge, Mass

Collins H M (1992a) 'Hubert Dreyfus, Forms of Life and a Simple Test For Machine Intelligence' Social Studies of Science Vol 22:726-739

Collins H M (1992b) 'Reproducing the Past: Three methods for Assembling A Cultural Inventory' Paper presented to the International Workshop on Replications of Historical Experiments in Physics, Oldenberg, August 24-26, 1992

Collins H M and Kusch M (in preparation) 'Two Kinds of Action: A Phenomonological Study'

Collins H M and Pinch T J (1993) 'The Golem: What everyone should know about science' Cambridge University Press, Cambridge

Collins H M and Yearley S (1992) 'Epistemological Chicken' in Pickering A (ed) 'Science as Practice and Culture' University of Chicago Press, Chicago

Cooley, M (1988) 'Creativity, Skill and Human Centered Systems' in Goranzon, B and Josefson, I (eds) Knowledge, Skill and Artificial Intelligence Berlin, Springer-Verlag.

Crean G P (1988) in Smith Kline and French laboratories publication 'Can GLADYS improve the management of dyspepsia?' The medicine group (UK)

Crean G P et al (1982) 'Ulcer-Like Dyspepsia' Scand J Gastroenterology Vol 17 (579):9-14

Critikon publication (I) 'Dinamap monitors. You know the name....' available from Critikon, Broadlands, Sunninghill, Ascot, Berkshire

Critikon publication (II) 'A Guide to Blood Pressure Measurement Techniques' available from Critikon, Broadlands, Sunninghill, Ascot, Berkshire

Davis M K (1982), 'Anaesthesia and the Neurosurgical Patient: Part I - Monitoring Needs' Journal of the American Association of Nurse Anaesthetists quoted in 'Survey of Studies involving the Dinamap Vital Signs Monitor' (1986) available from Critikon, Berkshire

Dawe V (1993) 'One-Third of juniors in dark on BP tests' Hospital Doctor 14 January

de Dombal T (1979) 'Computers and The Surgeon - A Matter of Decision' Surgery Annal (11):33-57

Department of Trade and Industry Publication (1990) 'Expert system opportunities. Complex Uncertain Decisions: Medical Diagnosis' London, HMSO

Doue J C and Vallance A G (Sept. 1985) 'Computer-Aided ECG Analysis' Hewlett Packard Journal:29-34

Dreyfus H L (1972, reprinted 1992) 'What Computers Can't Do' MIT Press, Cambridge Mass

Dreyfus H L (1987) 'Misrepresenting Human Intelligence' in Born R (ed) 'Artificial Intelligence: The Case Against' Croom Helm, Kent.

Dreyfus H L (1991) 'A Response to H M Collins', Discussion paper presented at 4S Conference, Boston, Mass, November 1991

Dreyfus H L (1992) 'Response to H M Collins. Artificial Experts, Social Knowledge and Intelligent Machines' Social Studies of Science Vol 22:717-726

Dreyfus H L and Dreyfus S E (1984) 'From Socrates to Expert Systems. The limits of Calculative Rationality' Technology in Society 6:217-233.

Dreyfus H L and Dreyfus S E (1986) 'Mind over Machine' The free press, NY.

Dreyfus H L and Dreyfus S E (1988) 'Making a Mind Versus Modelling the Brain. Artificial Intelligence at a Branchpoint' in Graubard S R (ed) 'The Artificial Intelligence Debate' MIT Press, Cambridge Mass

Duda R O and Shortliffe E H (1983) 'Expert Systems Research' in Science 220 (4594):261-268

Dunwoodie W M (Feb.1987) 'Computer aided diagnosis in dyspepsia' The Practitioner 231:162-168

Dylan, B (1963 - unreleased) 'Last Thoughts On Woody Guthrie' reproduced 1991 - 'The Bootleg Tapes, vol 1' Columbia Records

East T D et al (1990) 'Development of Computerised Critical Care Protocols - A strategy that Really Works' in 'Proceedings of the 14th annual SCAMC' - American Medical Informatics Association

East T D et al (1992) 'A strategy for the Development of Computerised Critical Care Decision Support Systems' Int J Clin Mon and Comput vol 8:263-269

Engelhardt H T (1974) 'Explanatory Models in Medicine: Facts, Theories and Values' Texas Reports on Biology and Medicine, Vol 32 (1):225-239

Epstein R H et al (1989) 'Evaluation of A Continuous Non-Invasive Blood Pressure Monitor in Obstetric Patients Undergoing Spinal Anaesthesia' J of Clin Monitoring Vol 5 (3), quoted in Survey of Studies Involving The Dinamap Vital Signs Monitor (1991), Available from Critikon, Berkshire

Evans R S (1991) 'The HELP system: A Review of Clinical Applications in Infectious Diseases and Antibiotic Use' MD Computing vol 8 (5):282-288

Evans R S et al (1985) 'Development of a Computerized Infectious Disease Monitor (CIDM)' Computers and Biomedical Research vol 18:103-113

Evans R S et al (1986) 'Computer Surveillance of Hospital-Acquired Infections and Antibiotic Use' JAMA Vol 256 (8):1007-1011

Evans W (1993) 'A Quantitative Approach to Understanding the Relationship Between Symbolic and Sub-Symbolic AI' Unpublished MSc project, University of Bath.

Feigenbaum E A and McCorduck P (1984) 'The Fifth Generation: Artificial Intelligence and Japan's Computer Challenge to the World' Pan Books, London

Fogel B S (1980) New Engl J Med 303 (22): 1307

Forster F K and Turney D (1986) 'Oscillometric Determination of Diastolic, Mean and Systolic Blood Pressure - A Numerical Model' Journal of Biomedical Engineering Vol 108:359-364

Forsyth R (1984) 'The Architecture of Expert Systems' in Forsyth R (ed)' Expert Systems: Principles and Case Studies' Chapman and Hall, London

Forsythe D (1992) 'Blaming the User in Medical Informatics: The Cultural Nature of Scientific Practice' Knowledge and Society Vol 9:95-111

Fox J and Alvey P (1983) 'Computer Assisted Medical Decision Making' BMJ Vol 287:742-746

Freisen R H and Lichtor J L (1981) 'Indirect Measurement of Blood Pressure in Neonates and Infants Utilising an Automatic Noninvasive Oscillometric Monitor' Anaesthesia and Analgesia (60) 10, quoted in 'Survey of Studies involving the Dinamap Vital Signs Monitor' (1986) available from Critikon, Berkshire

Gardner R M et al (1992) 'Editorial: Computing in the ICU: Is it feasible and practical?' International Journal of Clinical Monitoring and Computing vol 8:235-236

Gardner R M et al (1989) 'The Role of Smart Medical Systems in the Space Station' International Journal of Clinical Monitoring and Computing vol 6:91-98

Gattinoni L et al (1984) 'The Role of Total Static Lung Compliance in the Management of Severe ARDS Unresponsive to Conventional Treatment' Intensive Care Medicine vol 10:121-126

Gilbert G N and Mulkay M (1984) 'Opening Pandora's Box: A Sociological Analysis of Scientists' Discourse' Cambridge University Press, Cambridge

Ginzton L E and Laks M M (1984) 'Computer Aided ECG Interpretation' MD Computing Vol 1 (3):36:44

Gooding D C (1990) 'Experiment and The Making of Meaning: Human Agency in Scientific Observation and Experiment' Kluwar Academic Publishers, The Netherlands

Goranzon, B (1988) 'The Practice of the Use of Computers: A Paradoxical Encounter between different traditions of Knowledge' in Goranzon, B and Josefson, I (eds), 'Knowledge, Skill and Artificial Intelligence' Berlin, Springer-Verlag.

Gordon D R (1988) 'Clinical Science and Clinical Expertise: Changing Boundaries between Art and Science in Medicine' in Lock M and Gordon D R (eds) 'Biomedicine Examined' Kluwar Academic Publishers, Dordrecht, The Netherlands

Gorry G A (1973) 'Computer Assisted Clinical Decision Making' Meth Inform Med, 12 (1):45-51

Grene M (ed) (1969) 'Knowing and Being. Essays by Michael Polanyi' Routledge and Kegan Paul, London

Grundy B L et al (1981) 'Non-Invasive vs. Invasive Blood Pressure Measurements' Society of Critical Care Medicine original presentation, Washington DC. quoted in 'Survey of Studies involving the Dinamap Vital Signs Monitor' (1986) available from Critikon, Berkshire

Gullers, P (1988) 'Automation-Skill-Apprenticeship' in Goranzon, B and Josefson, I (eds) 'Knowledge, Skill and Artificial Intelligence' Berlin, Springer-Verlag.

Halfpenny P (1979) 'The Analysis of Qualitative Data' The Sociological Review Vol 27 (4):799-824

Hartland J (1990a) 'The Bottom Line' New Statesman/New Society, January 19 :26

Hartland J (1990b) 'Diagnosing the Normal: Diagnostic ECG Machines, Scope and Limitations' Cardiology in Practice Vol 8 (6):24-27

Hartland J (1993) 'The Use of 'Intelligent' Machines for Electrocardiograph Interpretation' in Button G (ed) 'Technology in Working Order', Routledge, London

Hatt G (1992) 'Disciplining the Body - The Production of Order in Medical Diagnosis' Unpublished MSc dissertation, University of Bath

Haugeland J (1985) 'Artificial Intelligence: The Very Idea' MIT Press, Cambridge, Mass

Health Devices Study (1987) 'Automatic Sphygmomanometers' ECRI Vol 15 (7) quoted in Survey of Studies Involving The Dinamap Vital Signs Monitor (1991) available from Critikon, Berkshire

Henderson S E et al (1990) 'Computerised Clinical Protocols in an Intensive Care Unit: How well are they followed?' Proceedings of the 14th annual SCAMC - American Medical Informatics Association

Henderson S E et al (1992) 'Performance of Computerised Protocols for the Management of Arterial Oxygenation in An Intensive Care Unit' Int J Clin Mon and Comput vol 8:271-280

Inoue T (1988) 'The Use of An Automated Blood Pressure Device - The Dinamap Monitor' The Clinical Report Vol 22 (14), quoted in Survey of Studies involving the Dinamap Vital Signs Monitor (1991), available from Critikon, Berkshire

Jamieson M J et al (1990) 'The Measurement of Blood Pressure: Sitting or Supine, Once or Twice?' Journal of Hypertension 8:6435-6440

Jones B (1983) 'The Division of Labour and the Distribution of Tacit Knowledge in the Automation of Metal Machining', Paper to IFAC conference, Germany, 7-9 November, 1983.

Jones B (1989) 'When Certainty Fails: Inside the Factory of The Future' in Wood S (ed) 'The Transformation of Work? Skill, Flexibility and the Labour Process' Unwin Hyman, London

Jones B (1990) 'Some Recent Features of Industrial Skill: Redefining its Social Construction' Paper to conference: 'Rediscovering Skill in Science, Technology and Medicine' Bath, September 1990.

Kingsland L C (1986) 'Artificial Intelligence: Promise and Reality' AAMSI Congress

Knill-Jones R (1986) 'A computer assisted diagnostic system for dyspepsia (GLADYS)' Lecture notes in medical informatics no.26

Knill-Jones R (1987) 'Diagnostic systems as an aid to clinical decision making' BMJ 295:1392-1396

Knill-Jones R, Dunwoodie W M, Crean G P (1985) 'A computer-assisted diagnostic decision system for dyspepsia' in Sheldon M, Brooke J and Rector A (eds) 'Decision making in general practice' Macmillan Press, Hampshire England

Koenig B A (1988) 'The Technological Imperative in Medical Practice: The social creation of a "Routine Treatment' in Lock M and Gordon D (eds) 'Biomedicine Examined' Kluwar Academic Publishers, Dordrecht, Netherlands

Kuhn T S (1962, reprinted 1972) 'The Structure of Scientific Revolutions' University Of Chicago Press, Chicago

Kuperman C J et al (1991) 'HELP - A Dynamic Hospital Information System' Springer Verlag, New York

Kusterer, K (1978) 'Know-How On The Job' Westview Press, Boulder, Colorado

Lack J et al (August 1989) 'An Assessment of Computerised Electrocardiograph Analysis' Cardiology in Practice Vol 7 (8):24/25

Larson R A et al (1989) ' Improved Perioperative Antibiotic Use and Reduced Surgical Wound Infections Through Use of Computer Decision Analysis' Infect Control Hosp Epidemiol vol 10 (7):316-320

Latour B and Johnson J (1988) 'Mixing Humans with Non Humans: Sociology of a Door Opener' Social Problems 35:298-310

Latour B and Woolgar S (1979) 'Laboratory Life: The Social Construction of Scientific Facts' Sage, Beverley Hills

Lilford R J and Thronton R D (1992) 'Decision Logic in Medical Practice', (The Milroy Lecture) Journal of the Royal College of Physicians Vol 26 (4):400-411

Lipscombe B R (1990) 'Minds, Machines and Medicine: An Epistemological Study of Computer Diagnosis' PhD Thesis, University of Bath

Macloed J (1979) 'Clinical Examination' Churchill Livingstone Longman Group, Edinburgh

Mackenzie D (1989) 'From Kwejelein to Armageddon? Testing and the Social Construction of Missile Accuracy' in Gooding D G, Pinch T J and Schaffer S (eds) 'The Uses of Experiment: Studies in the Natural Sciences' Cambridge University Press, Cambridge

Madsen E M et al (1991) 'Applications of Expert Computer Systems' Journal of the American Optometric Association Vol 62 (2):116-122

Marey E J (1879) 'Pression et Vitesse du Sang' Physiologie Experimentale Vol 2, Chapter VIII:307-343

Margolis J (1993) 'Wall to Wall Intellect' The Sunday Times (Section 9 - The Culture) 15 August, p. 6

Maxmen J S (1976) 'The Post Physician Era' John Wiley, NY

Maxmen J S (1987) 'Long Term Trends in Health Care: The Post Physician Era Reconsidered' in Detlef Schwefel (ed) 'Indicators and Trends in Health Care' Springer Verlag, Berlin

McCartney F J (1987) 'Diagnostic logic' BMJ 295:1325-1331

Medawar P B (1967) 'The Art of The Soluble: Creativity and Originality in Science' Methuen, London

Meehl P E (1954) 'Clinical Vs Statistical Prediction' University of Minnesota Press, Minneapolis

Merton R K (1952) 'The Neglect of the Sociology of Science' chapter 10 in Storer N W (ed) (1972) 'The Sociology of Science: theoretical and empirical investigations' University of Chicago Press, Chicago

Miller B F and Keane C B (1978) 'The Encyclopedia and Dictionary of Medicine, Nursing and Allied Health' W B Saunders Co, London

Miller R A et al (1982) 'Internist 1, an experimental computer based diagnostic consultant for general internal medicine' N Engl J Med, 307:468-76

Miller RA et al (1986) 'The INTERNIST-1/QUICK MEDICAL REFERENCE Project - Status Report' The Western Journal of Medicine Vol 145 (6):816-821

Minsky M (1974) 'A Framework for Representing Knowledge', MIT Lab memo 306, excerpts in Haugeland J (ed) (1981) 'Mind Design' MIT Press, Cambridge, Mass

Minsky M and Papert S (1969) 'Perceptrons: An Introduction to Computational Geometry' MIT Press, Cambridge Mass

Morris A H (1992) 'Protocols, ECCO2R and the Evaluation of New ARDS Therapy' Japanese Journal of Intensive Care Medicine vol 16:61-63

Morris A H (in press) 'Paradigms in Management' in Pinsky M R, Dhounaut J F A (eds) 'Pathophysiologic Foundations of Critical Care Medicine' Williams and Wilkins, Baltimore, MD

Morris A H et al (1989) 'A Protocol Controlled Clinical Trial of a New Three Step Therapy including ECCO2R for Adults with ARDS' in Gille J P (ed) 'Neonatal and Adult Respiratory Failure: mechanisms and treatments' Elsevir, Paris

Morris A H et al (1990) 'Extracorporeal CO2 Removal Therapy for Adult Respiratory Distress Syndrome Patients' Respiratory Care vol 35 (3):224-231

Morris A H et al (1992) 'Final Report: A Computerised Protocol Control Randomized Clinical Trial of New Therapy that includes LFPPV-ECCO2R' in 'Proceedings of 6th Annual Congress on Intensive care Medicine' 26-31 October 1992, Barcelona

Mulkay M (1985) 'The Word and The World: Explorations in the form of Sociological Analysis' George Allen and Unwin, London

Mulkay M, Ashmore M and Pinch T (1987) 'Measuring the Quality of LIfe: A Sociological Invention Concerning the Application of Economics to Health Care' Sociology vol 21 (4):541-564

Mulkay M, Pinch T and Ashmore (1987) 'Colonizing the Mind: Dilemas in the Application of Social Science' Social Studies of Science vol 17:231-256

Newman N et al (1988) 'Microcomputer-aided diagnosis of jaundice (SOLUBILE)', in Vicary F R (ed) 'Computers in Gastroenterology' Springer-Verlag, London

Nii (1984) in Feigenbaum E A and McCorduck P (1984) 'The Fifth Generation: Artificial Intelligence and Japan's Computer Challenge to the World' Pan Books, London

Noble D (1984) 'Forces of Production' Alfred A Knopf, New York

O'Brien E and O'Malley K (1990) 'Twenty-Four Hour Ambulatory Blood Pressure Monitoring: A Review of Validation Data' Journal of Hypertension, Vol 8, Supplement 16:S11-S16

Olazaran M (1991) 'A Historical Sociology of Neural Network Research' PhD Thesis, University of Edinburgh

Park M K and Menerd S M (1986) 'Accuracy of Blood Pressure Measurement by the DINAMAP in Infants and Children' Journal of Pediatrics quoted in 'Survey of Studies involving the Dinamap Vital Signs Monitor' (1986) available from Critikon, Berkshire

Pasveer B (forthcoming) 'Coding and Decoding: The skill of reading Medical X-Ray Photographs'

Pelosi A J and Lewis G (1989) 'The Computer Will See You Now' BMJ Vol 299:138-139

Perloff J K (1982) 'Physical Examination of The Heart and Circulation' W B Saunders, London

Pestotnik S L et al (1990) 'Therapeutic Antibiotic Monitoring: Surveillance Using a Computerised Expert System' The American Journal of Medicine vol 88:43-48

Petrie J C et al (1986) 'British Hypertension Society Recommendations on Blood Pressure Measurement' BMJ vol 293:611-615

Pinch T (1993) 'Generations of SSK' Social Studies of Science Vol 23:363-375

Polanyi M (1962) 'Personal Knowledge' Routledge and Kegan Paul, London

Pothoff P et al (1988) 'Expert Systems in Medicine: Possible Future Effects' Intl. J of Technology Assessment in Health Care Vol 4:121-133

Priest C (1989) Hewlett-Packard Memo, Personal Correspondence

Pryor T A et al (1983) 'The HELP System', Journal of Medical Systems vol 7(2):87-102

Ramsay M (1979) 'Noninvasive Automatic Determination of Mean Arterial Pressure' Medical and Biology Eng. and Computer (17) 11:1, quoted in 'Survey of Studies involving the Dinamap Vital Signs Monitor' (1986) available from Critikon, Berkshire

Rector A L (1984) 'Knowledge Based Systems in Medicine: A Review' in Expert Systems: State of the Art Report 12 (7), Pergamon Infotech, Maidenhead

Resnick M (Aug. 1982) 'Artificial Intelligence: The Second Computer Age Begins' Business Week 8 August 1982:66-75

Rolfe H (1986) 'Skill, Deskilling and New Technology' New Technology, Work and Employment vol 1 (1):37-49.

Rolfe H (1990) 'In the Name of Progres? Skills and Attitudes Towards Technological Change', New Technology, Work and Employment, vol 5 (20):107-121

Rosenblatt F (1962) 'Principles of Neurodynamics: Perceptrons and The Theory of Brain Mechanisms' Spartan Books, NY

Rudwick M (1985) 'The Great Devonian Controversy' University of Chicago Press, Chicago

Schank R and Abelson R P (1977) 'Scripts, Plans, Goals and Understanding' Hillsdale, New Jersey

Schank R and Riesbeck C (1981) 'Inside Computer Understanding' Hillsdale, New Jersey

Schwartz W B (1970) 'Medicine and the Computer, the Promise and Problems of Change' New Engl J Med, 223 (23):1257-64

Schwartz W B, Patil R S, Szolovits P (1987) 'Artificial Intelligence in Medicine. Where do we stand?' N Engl J Med, 316 (11):685-688

Seaman D (1985) 'Should you trust automatic Blood Pressure Monitors? Nursing '85 quoted in 'Survey of Studies involving the Dinamap Vital Signs Monitor' (1986) available from Critikon, Berkshire

Searle J (1981) 'Minds, Brains and Programs' in Hofstadter D R and Dennett D C (eds) 'The Minds Eye' The Harvester Press, Brighton

Searle J (1984) 'Minds, Brains and Science' BBC publications, London

Short D (1976) 'The Diastolic Dilemma' BMJ Vol 2:685-686

Shortliffe E H, Buchanan B G and Fiegenbaum E A (1979) 'Knowledge Engineering for Medical Decision Making: A Review of Computer Based Clinical Decision Aids' in Clancey W J and Shortliffe E H (eds) Readings in Medical AI - The First Decade Addison-Wesley, Reading, MA

Slezak P (1992) 'Essay Review - Artificial Experts' Social studies of Science 22:175-201

Social Studies of Science - Special Issue, vol 11, no 1 (1981) 'Knowledge and Controversy: Studies of Modern Natural Science' edited by H M Collins

Sorgaard P (1991) 'Evaluating Expert System Prototypes' AI and Society (5):3-17

Spiegelhalter D J and Knill-Jones R (1984) 'Statistical and Knowledge-based approaches to clinical decision-support systems, with an application in Gastroenterology' J R Statist. Soc A 147:35-77

Storer N W (ed) (1972), 'The Sociology of Science: theoretical and empirical investigations' University of Chicago Press, Chicago

Suchman L A (1987) 'Plans and Situated Actions: the problem of Human Machine Communication' Cambridge University Press, Cambridge

Sutton G C (1989a) 'Computer Aided Diagnosis: A Review' Br J Surg Vol 76:82-85

Sutton G C (1989b) 'How Accurate is Computer Aided Diagnosis?' The Lancet Oct 14:905-908

Szolovits P ed (1982) 'Artificial Intelligence and Medicine' in Szolovits P (ed) 'Artificial Intelligence in Medicine' Westview Press, Boulder, Colorado

Szolovits P and Pauker S G (1979) 'Computers and Clinical Decision Making: Whether, How and For Whom' Proceedings IEEE Vol 67 (9)

Szolovits P, Patil R S and Schwartz W B (1988) 'Artificial Intelligence in Medical Diagnosis' Annals of Internal Medicine (108):80-87

Szolovits P and Pauker S G (1993) 'Categorical and Probabilistic Reasoning in Medicine Revisited' Artificial Intelligence (59):167-180

Taylor T R (1990) 'The Computer and Clinical Decision Support Systems in Primary Care' The Journal Of Family Practice Vol 30 (2):137-140

Taylor, F W (1947) 'Principles of Scientific Management' Harper and Row, NY

Thornton J G and Lilford R J (1992) 'Decision Analysis in Medicine' BMJ Vol 304:1099-1103

Turing A M (1936) 'On Computable Numbers, with an application to the Entscheidungs Problem' Proceedings of the London Mathematical Society 2nd Series Vol 42:230-265 and Vol 43:544-546 (1937)

Van Der Lei J et al (1991) 'Comparison of Computer-Aided and Human Review of General Practitioners' Management of Hypertension' The Lancet Vol 338:1504-1508

Van Rijsbergen C J (1985) 'Turing and the Origins of Digital Computers' Aslib Proceedings, Vol 37 no 6/7:281-285

Wallace C J (1993) Personal correspondence - fax February 18, 1993

Warner H R (1978) 'First the Electrocardiogram - Then What?' The AM J of Cardiology vol 41:115-118

Warner H R (1980) 'HELP - An approach to Hospital Wide Artificial Intelligence' in Statland B F and Bauer S (eds) 'Computer Assisted Decision Making Using Clinical and Para Clinical (Laboratory) Data' Mediad Inc, USA

Weizenbaum J (1976) 'Computer power and Human Reason' The Chaucer Press, Suffolk

Winograd T (1972) 'Understanding Natural Language' emic Press, NY

Winograd T and Flores F (1986) 'Understanding Computers and Cognition' Ablex Corporation, Norwood, NJ

Winston P H and Prendergast K A (1984) 'The AI Business' MIT Press, Cambridge Mass

Wittgenstein L (1958) 'Philosophical Investigations' Blackwell, Oxford

Wolpert L (1992) 'The Unnatural Nature of Science' Faber, London

Woolgar S (1985) 'Why Not A Sociology of Machines? The case of Sociology and Artificial Intelligence' Sociology (19) 4:557-572

Woolgar S (1988) 'Science - The Very Idea' Tavistock, London

Woolgar S (1992) 'Some Remarks about Positionism: A Reply to Collins and Yearley' in Pickering A (ed) 'Science as Practice and Culture' University of Chicago Press, Chicago

Woolgar S and Ashmore M (1988) 'Knowledge and Reflexivity' Sage, London

Wright P and Treacher A eds (1982) The Problem of Medical Knowledge: Examining the Social Construction of Medicine Edinburgh University Press, Edinburgh

Wright P K and Bourne D A (1988) 'Manufacturing Intelligence' Addison Wesley, Mass

Wyatt J (1991) 'Computer Based Knowledge Systems' The Lancet Vol 338:1431-1436

Yu V L et al (1979) 'Antimicrobial Selection by a Computer: A Blinded Evaluation By Infectious Diseases Experts' JAMA Sept 21, Vol 242 (12):1279-1282

Ziman J (1984) 'An Introduction to Science Studies - the philosophical and social aspects of science and technology' Cambridge University Press, Cambridge, England

APPENDIX 1

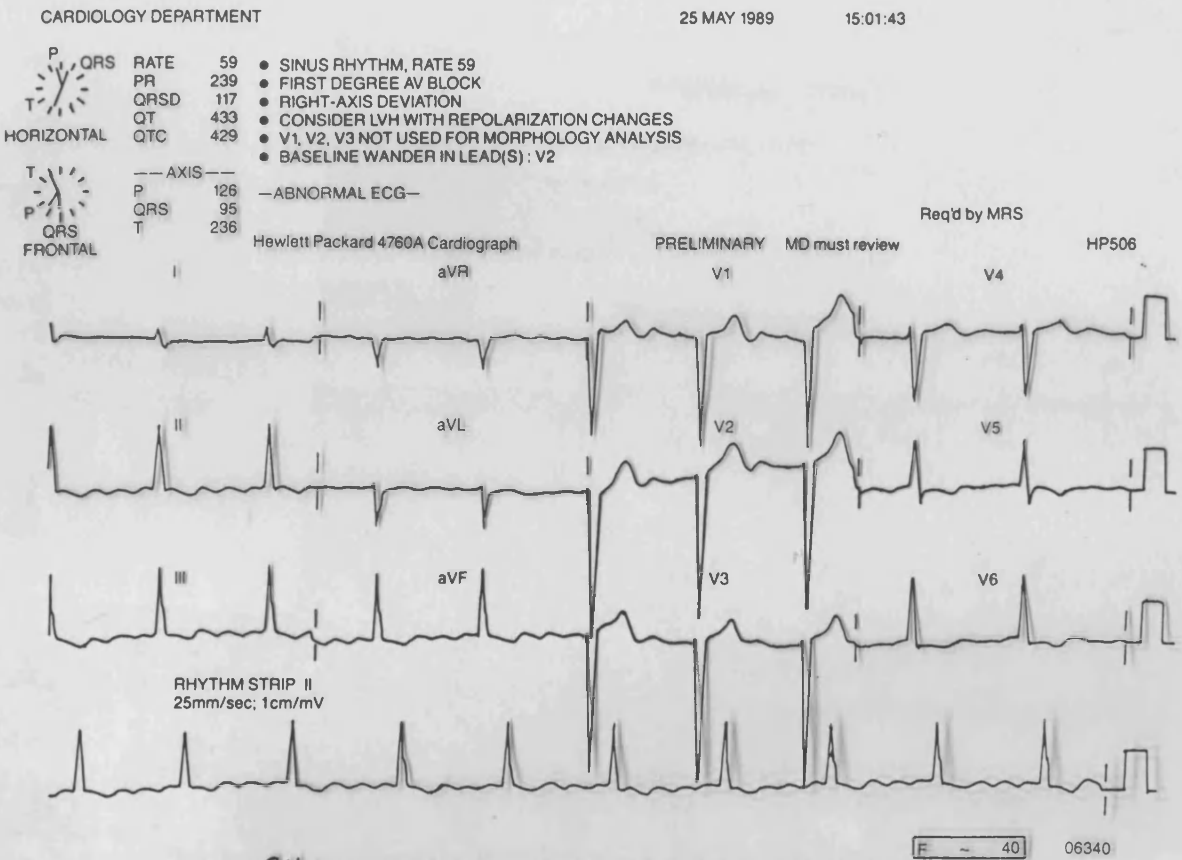


Figure 8·4 Interpreted as: sinus rhythm with first degree AV block – 15:01 hrs

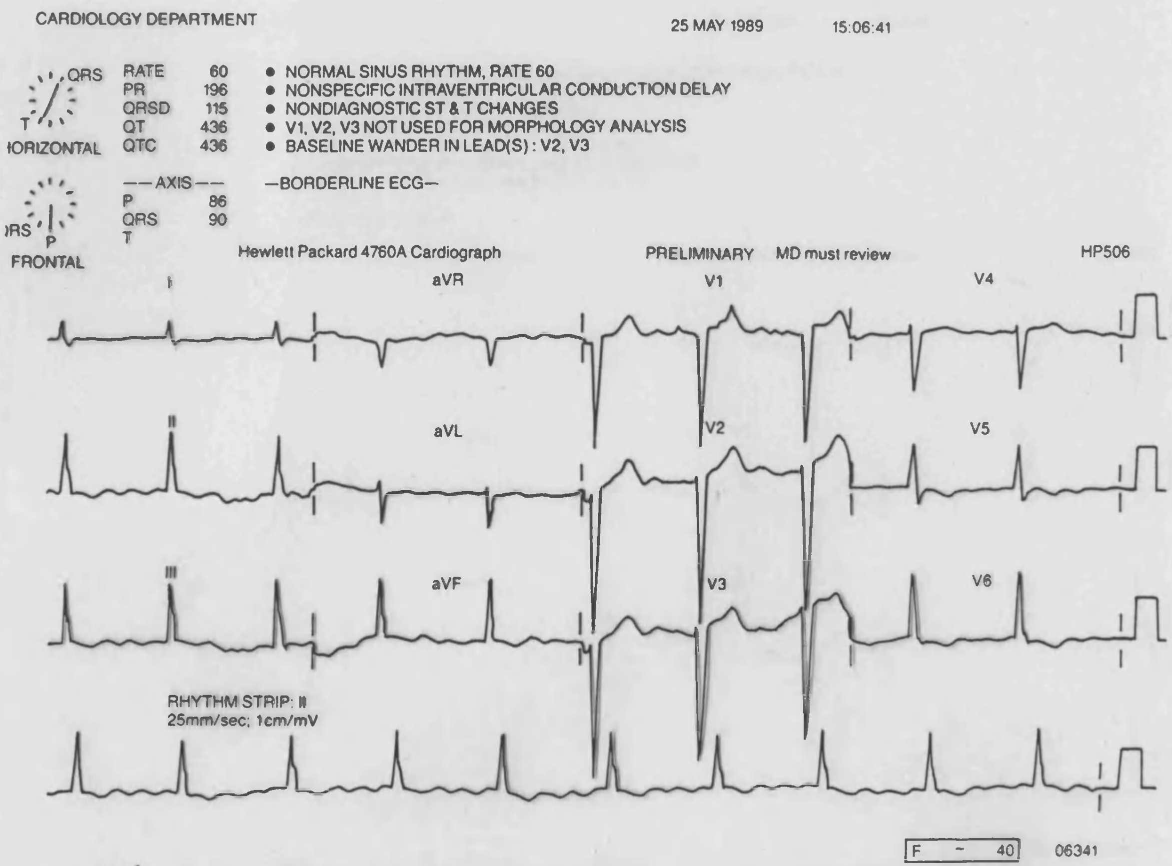


Figure 8.5 Interpreted as: sinus rhythm with non-specific intraventricular conduction delay – 15:06 hrs

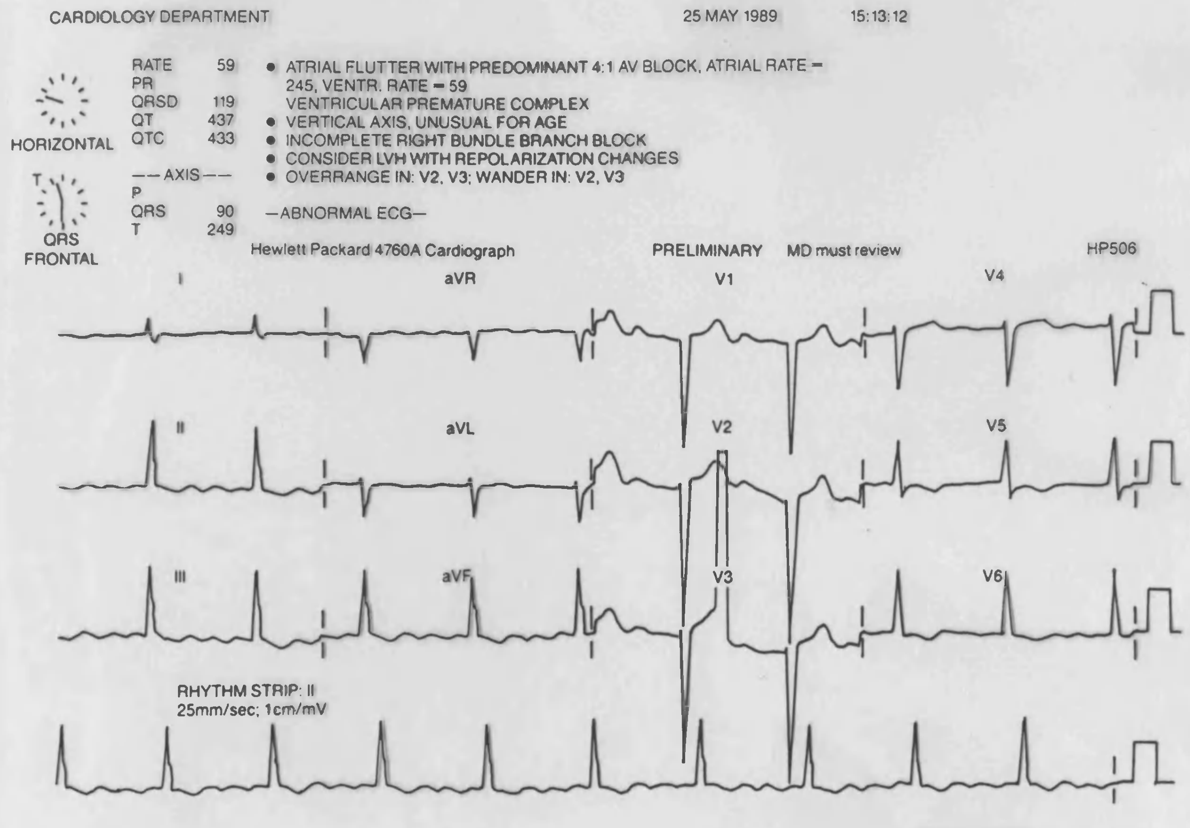


Figure 8·b Interpreted as: atrial flutter with predominant 4:1 AV block – 15:13 hrs