OBSERVATIONS OF MEDICAL PROFESSIONALS’ INTERACTIONS WITH AN INTELLIGENT TUTORING SYSTEM

BY

DAVID C. WILLIAMS

A THESIS PRESENTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH, IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF ARTS

DEPARTMENT OF EDUCATIONAL PSYCHOLOGY AND COUNSELLING
MCGILL UNIVERSITY, MONTREAL
MAY 1990

© DAVID CHARLES WILLIAMS, 1990
With Gratitude and Affection to my loving parents

Joyce Eileen and Sydney Alexander Williams
ACKNOWLEDGEMENTS

I am most grateful for and to the many people for their help with regards to this dissertation. Although there is not enough space here to sufficiently thank all those who should be thanked, the author wishes to express his gratitude to Anne Løwen, and all the physicians who volunteered their time to participate in this study, who were so very gracious and helpful. I shall never forget their kindness and their patience. The generous help of Dr. Robyn Tamblyn Dr. Albert Larbisseau, Dr. Tanoni, and Dr. Ruth Horn, who acted as expert consultants, can only be described as above and beyond the call of duty; without their input my assessments would have been impossible. I would also like to acknowledge the invaluable support of my parents Joyce and Sydney Williams; who shared my ups and downs, and my many friends who supported me. My supervisor, with whom I could bounce ideas off of, Dr. Guy Joseph Groen, provided inspiration, and encouragement throughout the study and for this I am most grateful. I would like to thank Dr. Vimla Patel, for supplying me with resources and urging me on to further academic work. I wish to specially thank David Kaufman, Jose Arocha, and Lorence Coughlin for their invaluable friendship and time. I also wish to thank Aldo Braccio, Alejandro Larrivee, Carole Levesque, Céline Pêreira, Denis Ratté, Dr. John Macnamara, Dr. Nicole Domingue, Dr. Rod MacDonald, Dr. Socrates Rapagna, Dr. W. Dale Dauphinee, Dr. Angel Rodriguez, Ezio Radeschi & Elena Pedicelli, Gustavo Tanoni, Daniel Simon, Michael Kunst, Dr. Michael Dillinger, Paul Cormier, Riccardo Magini & Louise Zampini, Richard Brisebois Malathi Sivaramakrishnan and all of my friends. A special thanks to Patricia Grafton, John St. G. McCabe, Paul Waters and Luis San Salvador. I most certainly would have given up a long time ago if not for them. It is an honour and indeed a very special and distinctive privilege to be among such company. I shall never be able to repay them.
ABSTRACT

Intelligent tutoring systems (ITS) are expert systems united with computer-aided instruction. The psychological issue of human-computer interfacing combines aspects of education, cognitive science, human performance and psycho-sociolinguistics. This study presented a situation in which physicians used their reasoning to solve a computer-simulated medical case, embedded in the NEOMYCIN ITS. Experiments were designed to assess how their anthropomorphisation of the systems affected their medical reasoning in a complex ill-defined problem-solving domain. The study examines the subjects' interpretation of textual case materials, specifically their ascription of meaning and intelligibility to the form and usage of natural language. The results indicate that these factors affect their interpretation, not only of case materials, but also of their evaluation of the program's medical reasoning. This has implications for the interactive man-machine interface and its relationship to interpersonal communication is discussed.
RÉSUMÉ

Les systèmes intelligemment assisté (STI) sont des systèmes experts liés avec un enseignement assisté d'un ordinateur. Le côté psychologique de l'interface homme-ordinateur combine des aspects de l'éducation, des sciences cognitives, de la performance humaine et de la psycho-sociolinguistiques. Cette étude a présenté une situation dans laquelle les médecins qui y travaillaient ont utilisés leur raisonnement pour solutionner un cas médical simulé par un ordinateur, est inclu dans le système NEOMYCIN. Des expériences ont été conçues pour évaluer comment l'antropomophisation des systèmes a affecté leur raisonnement médical dans un domain complexe, mal défini et concernant la solution de problèmes. L'étude évalue l'interprétation par les sujets de matériaux d'un cas textuel, en particulier leur attribution du sens et de l'interprétation à la forme et à l'usage de la langue naturelle. Les résultats démontrent que ces facteurs affectent leur compréhension, non seulement des matériaux du cas, mais aussi de leur évaluation du raisonnement médical du programme. Ceci a des implications pour l'interface interactive homme-machine et sa relations avec la communication interpersonnelle est discutée.
TABLE OF CONTENTS

Acknowledgements ................................................................. i
Abstract ......................................................................................... ii
Résumé .......................................................................................... ii
List of Tables .................................................................................... vi
List of Figures ................................................................................... vii
List of Appendices ........................................................................ ix

Chapter I ............................................................................................... 1
INTRODUCTION ..................................................................................... 1

CHAPTER II ......................................................................................... 6
REVIEW OF THE LITERATURE ............................................................ 6
The Focus ............................................................................................ 9
Communication ................................................................................... 12
The Knowledge & the Skill .................................................................. 13
What is NEOMYCIN? ......................................................................... 17
MYCIN ............................................................................................... 21
Natural Language ................................................................................ 22
Modelling ........................................................................................... 26
Pragmatics ......................................................................................... 29
Taking for Granted What We Know ................................................... 29
Natural Language Environments ....................................................... 35
Seeing Things ................................................................................... 41
CONSOLIDATION ................................................................................ 42

CHAPTER III ......................................................................................... 46
PURPOSE OF THE STUDY ................................................................. 46
Questions to be Investigated in this Study .......................................... 48
Foundational Issues ......................................................................... 48
Hypotheses for this Study ................................................................. 51

CHAPTER IV ......................................................................................... 54
BACKGROUND: THE MEDICAL CASE ................................................ 54
Differential Diagnoses ....................................................................... 54
The Task ............................................................................................ 55
The Problem .................................................................................... 56
THE CASE DEBRA ............................................................................... 56
The Diagnosis ................................................................................ 59
Meningitis ....................................................................................... 61

CHAPTER V ......................................................................................... 63
METHODOLOGY OF THE STUDY ...................................................... 63
METHOD ............................................................................................ 63
EXPERIMENT I ..................................................................................... 64
Subjects and Apparatus .................................................................... 65
Preparation ...................................................................................... 65
Procedure ....................................................................................... 66
The Interaction ................................................................................ 69

EXPERIMENT II .................................................................................... 70
EXPERIMENT IIA ................................................................................. 70
Subjects and Equipment ................................................................. 70
LIST OF TABLES

<table>
<thead>
<tr>
<th>TITLE</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 1. Findings in Patient Case Debra</td>
<td>58</td>
</tr>
<tr>
<td>Table 2. Short Instruction</td>
<td>66</td>
</tr>
<tr>
<td>Table 3. Presenting Complaint</td>
<td>71</td>
</tr>
<tr>
<td>Table 4. Overview of Subjects</td>
<td>90</td>
</tr>
<tr>
<td>Table 5. Initial Diagnosis upon Presentation of Presenting Complaint</td>
<td>144</td>
</tr>
</tbody>
</table>
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>TITLE</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1. Reason for Differences between Expert &amp; Expert System</td>
<td>69</td>
</tr>
<tr>
<td>Figure 2. Facets of Experimental probes with respect to Subjects' responses to Findings presented by NEOMYCIN</td>
<td>76</td>
</tr>
<tr>
<td>Figure 3. Facets of Experimental probes: Subjects' responses to questions and explanations presented by NEOMYCIN</td>
<td>77</td>
</tr>
<tr>
<td>Figure 4. Types of rejoinders</td>
<td>82</td>
</tr>
<tr>
<td>Figure 5a. Schemata of Clinical Reasoning of Gerontologist. Subject #4</td>
<td>119</td>
</tr>
<tr>
<td>Figure 5b. Schemata of Clinical Reasoning of Gerontologist. Subject #4 (continued)</td>
<td>121</td>
</tr>
<tr>
<td>Figure 5c. Schemata of Clinical Reasoning of Gerontologist. Subject #4 (continued)</td>
<td>123</td>
</tr>
<tr>
<td>Figure 5d. Schemata of Clinical Reasoning of Gerontologist. Subject #4 (continued)</td>
<td>126</td>
</tr>
<tr>
<td>Figure 5e. Schemata of Clinical Reasoning of Gerontologist. Subject #4 (continued)</td>
<td>130</td>
</tr>
<tr>
<td>Figure 5f. Schemata of Clinical Reasoning of Gerontologist. Subject #4 (continued)</td>
<td>131</td>
</tr>
<tr>
<td>Figure 6. Network Schemata of Clinical Reasoning. Subject #4, Gerontologist</td>
<td>133</td>
</tr>
<tr>
<td>Figure 7. Network Schemata of Clinical Reasoning. Subject #1 Gerontologist</td>
<td>134</td>
</tr>
<tr>
<td>Figure 8. Network Schemata of Clinical Reasoning. Subject #2, Resident</td>
<td>136</td>
</tr>
<tr>
<td>Figure 9. Network Schemata of Clinical Reasoning. Subject #3 Resident</td>
<td>137</td>
</tr>
<tr>
<td>Figure 10. Network Schemata of Clinical Reasoning. Subject #6, Internist</td>
<td>138</td>
</tr>
<tr>
<td>Figure 11. Network Schemata of Clinical Reasoning. Subject #7, Resident</td>
<td>139</td>
</tr>
</tbody>
</table>
Figure 12a. 
Reasoning Considerations of Subject #1 

Figure 12b. 
Reasoning Considerations of Subject #1 (continued)

Figure 13. 
Schemata of Clinical Reasoning of Gerontologist. Subject #5
APPENDICES

<table>
<thead>
<tr>
<th>TITLE</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Appendix A</td>
<td>185</td>
</tr>
<tr>
<td>Patient Case 1000 (Debra)</td>
<td></td>
</tr>
<tr>
<td>Appendix B</td>
<td>188</td>
</tr>
<tr>
<td>Detailed Instructions and Experimental Protocol</td>
<td></td>
</tr>
</tbody>
</table>
CHAPTER 1

INTRODUCTION

With the advent of computer-based intelligent tutoring systems (ITS), it is of interest to educational psychologists to observe how computer users function in these complex microworld environments — a computer-based environment which a student can explore and possibly extend (Groen, 1985). As these systems are intended to create a problem-solving situation which is based, as close as possible, on the real-life knowledge of the user, it would not be surprising that people would use their real-world knowledge and skill in order to solve problems in these simulations. In so being, a simulation of reality, these systems cue the subject to think as they learn, rather than merely providing them with information.

Information from the environment can be perceived in basically different ways by human observers (Morris, 1971; Eco, 1979). One’s percept is the interpretation of one’s sensations, but the interpretation is not perceptually given (and the intelligence knows things which we cannot perceive). If this is true for our perception of reality, what about situations presented on computer technology? To examine our understanding of computer displays we need to know what things influence our cognition?

Moreover, with respect to language, when the interface is conversationally interactive (i.e., the subjects are communicating with machine and, vice versa) it is inevitable that aspects of human social action will affect user perception of the content of the problem-solving simulation. People may interact with the machine as though it were an intentional person. That is to say, that the users’ knowledge of language use, as an aspect of cooperative communication between a producer and a receiver, will play a part in their interpretation of the material presented by the programme— with speakers being more likely to attain their goals when they act on true beliefs (Putnam, 1983).
part in their interpretation of the material presented by the programme— with speakers being more likely to attain their goals when they act on true beliefs (Putnam, 1983).

...people's ability to understand language as it is used depends on their ability to go beyond what is literally asserted, to infer or otherwise arrive at assumptions, presuppositions, and unstated conclusions underlying the statements that actually occur, either as utterances or inscriptions. It must be assumed therefore, that anyone able to use language for conversational purposes must be a very skillful logician in order to supply all and only correct implicatures and inferences required for understanding, and to do so at the rates required for ordinary social discourse (Miller & Glucksberg, 1988; p. 464).

Language is a form of cooperative activity. Practical situations arise whenever people communicate pertaining to what one is expecting from one's interlocutor and what one feels he or she is expecting in return. The investigation of this human/computer discourse necessitates that one addresses certain issues, many of which are semiotic in nature.

Semiotics describes that which is about signs, or more specifically, signalling systems (Lyons, 1977; Levinson, 1983). A concept that picks out an object in the environment is associated with a sign (in this case, a word). What one thus has in mind becomes the intended meaning of that sign (Putnam, 1988). We get to the real world because we relate to the experience of the speaker. Language is referred to, is applied to, and symbolizes our experience. The real world is neither shaped by our minds nor are our minds rigidly shaped by language. This has been shown in incidences of linguistic relativity between languages, like porte in French individuates more objects than door in English. Putnam (1975) proposes a division of labour between our nonlinguistic knowledge and the concepts behind our words. For example, "water" in everyday language need not be specified as to distinguish it from things that look like it, but
when talking to a chemist the conceptualization for the sign is deeper. Speakers know who their intended audiences are, but the term still individuates the same object.

Language is a type of signalling system in which a producer intends to convey a message and performs an act to communicate that message. That agent sends a signal, which has a form and a convention of usage, in order that the receiver may interpret that signal's form in terms of the usage conventions, as well as other contextual indices. What the signal encodes is cognitive content (some meaning). However, what the interlocutor might unconsciously embellish the sign with (extrasemantic message), might even transcend the communicator's intent. A simple statement might produce different reactions in the recipient if it were thought that a statement was said in anger, or with urgency, or concern. Prosody is one ingredient which one uses in order to determine these differences, which, understandably, are not as ostensibly present in written text, as would be presented on a computer screen. In addition, the context in which the message is received affects how it is interpreted. The signaler might not want to be interpreted in such a way, and an interpreter might misread the signal. This is because people attribute purpose, order, and meaning to all entities which they feel to be intelligible. Communication is no exception. Even if this exchange is between a computer programme, which functions as a surrogate of a programmer, and a user.

These communicative considerations fall under the auspices of *pragmatics* (Levinson, 1983). By linguistic pragmatics, I refer to the field of linguistics whose goal it is to discover the principles by which hearers or readers determine what something was really intended to mean (Prince, 1982).

In one's encounter with reality four basic questions are asked:

What is this?
What is it for?
What is it made of?
Who made it?
They form our common sense intuition (reality as it appears in the senses) about the world. Pragmatics assumes such consideration are also made about language exchange. For instance, Grice, (1975) posits a Maxim of Quantity which states that natural-language users assume that their interlocutors do not ask for information which they already have or can plausibly infer. Therefore, if information is redundant (in a communication), one assumes that the other party is being cooperative and thus one will decide that the writer/speaker is trying to convey something other than the literal meaning. That is, certain assumptions can be made about the meaning of a lexical string beyond its propositional content.

The research which follows constitutes the findings of a descriptive exploratory study about the nature of interpretation with respect to human/computer interaction in an intelligent tutoring system.

The next chapter: Chapter II discusses discrete issues related to providing the reader with background on elements of computer-based instruction, expert behaviour, and linguistic pragmatics. The chapter provides a review of the relevant literature, and serves to formulate and establish the basis for the construction of this study. It argues that if knowledgeable users are encouraged to explain their actions and to critique the "actions" of the programme, their intuitive understanding of clinical problem as well as their conversational considerations will become apparent in the experimental subjects' think aloud protocols.

Chapter III states the purpose and rationale of this study, which is to investigate interactive diagnostic sessions between medical professionals and computerised tutoring system. On this tutoring system, GUIDON (attached to the expert system programme NEOMYCIN), there is a simulated medical case. A list of hypotheses has been entertained in this chapter; the focus of the dissertation is to examine and to assess how certain linguistic pragmatics will affect the doctor's medical reasoning. Reasoning within will refer to the non-intuitive discursive process of running along a
path of thought or cognitive investigation. Johnson-Laird (1988) points out that we are not aware of how we reason; we observe only the consequences in our conscious thought.

Chapter IV illustrates the nature of the medical reasoning task, with direct reference to the medical case encountered in this study and its diagnosis. The chapter describes the problem-space of the findings and diagnostic hypotheses of the medical case used in this study. It includes a discussion of normal physiology, as well as the potential disease and disorder hypotheses which one could arrive at erroneously.

Chapter V contains a description of the equipment and methodologies employed in this study, as well as the methods of analyses. The method consists of a combination of think aloud protocols and specific probing for information.

Chapter VI presents the results of the study and discusses the significance of what the results of this work entail.

Chapter VII draws conclusions as well as implications for further research.
CHAPTER II

REVIEW OF THE LITERATURE

This chapter presents a review of issues related to user understanding in computerized intelligent tutoring system interfacing with respect to linguistic pragmatics. It also deals with aspects of human action and intentionality that affect interpretation and meaning.

As computers seem capable of carrying out routine and even intricate tasks repeatedly and consistently, they have been considered to be beneficial in education as interactive teaching devices (Sleeman & Brown, 1982; Shore, 1973; Mayer, 1975; Groen, 1984; Kearsley, 1987; Lawler & Yazdani, 1987). Computer-based instruction has been implemented in situations where a number of students need to perform 'drill and practice' exercises. The use of computerized teaching machines may, and indeed have, freed instructors from mundane tasks such as giving drill and practice. This gives human instructors more time to be able to do more 'reactive' teaching, responding to the on-going needs of a class. Computer programmes often can also be thought of as tutors in their own right (Clancey, 1987b; Sleeman & Brown, 1982). Many have an interesting interface designed to be varied, novel, and entertaining so that the student is more motivated to learn. This is especially applicable in repetitive tasks which require keeping the student 'on-task' for some time.

The use of technology in pedagogy, in combination with other procedures, and not substituting for personal attention, as in apprenticeship training, may be a useful way to help students acquire skills and techniques, in a particular domain. One aspect of education in which this technology can play a part, is to adapt group learning methods and techniques to the individual (e.g., the case method in explanation-based learning to concretize examples and heuristics). The programme used in this thesis is amenable to this type of usage. Lobowitz, (1986) describes explanation-based learning
on computers as a method to enhance the abstraction students make from an encounter with information in a single instance. He proposes developing a computer programme which builds up a series of explanations upon elements of a problem. This process is to show how the problem elements, relate to each other, showing the underlying method to solve the problem. The programme then generalizes the properties of the components of the instance as long as the explanations remains valuable. The student thus learns, not only about the instance, but about a class of instances in different circumstances.

Computer assisted instruction (CAI) can also provide simulations to a learner via exploitable ‘discrete’ or familiar ‘objects’ presented on a computer screen. This is in contrast with a system only designed to present static materials for simple rote learning and repetition. Simple learning, as higher learning, is acquired through experience and practice. However, a great many problem-solving tasks which are encountered in higher learning are not ‘well-structured’ (Greeno & Simon, 1986). An electronic tutor for ill-structured problems has to do more than inflexibly represent a few test exercises (Chi, Glaser, & Rees, 1982). Because ill-structured problems have first conditions, goals and essential parameters which are unexaminable and obscure at the inception of the problem-solving procedure (Simon, 1978) – there is no single best way to solve the problem. Because of the diversity of correct problem-solving paths, the programme would have to be resilient enough to deal with the many ways in which such problems can be authentically approached and solved by many different individuals. The programme would have to be an ‘expert’ in the area it teaches, with a tutoring programme to survey the needs of the student intelligently by monitoring the problem-solving task as well as the student’s operations. In addition, it would need to answer student questions that are related to the current learning material. The agenda required for teaching in more complicated domains seem to require ‘smart’ machines and ‘smart’ programmes. Inaccuracy in the evaluation of the needs of the user would be
pedagogically counter-productive. The teaching procedures would have to be either like a live teacher, or like an apparatus which exploits the attributes that computers and programmes have that human beings do not (Winograd & Flores, 1986).

Engineers and computer scientists are grappling with the implementation of artificial intelligence (AI) programmes. They have not succeeded attaining human-type intelligence on a machine. These are programmes which can 'make decisions', or model some sort of reasoning. Some of these programmes are not directly geared at modelling reasoning, but rather at performing specific tasks at or above the level of an expert in that field. These are called 'expert systems' (Hayes-Roth, Waterman, & Lenat, 1983).

There are attempts to bring this onslaught of computer and computational technology to address issues in the cognitive sciences. The cognitive sciences involve the fields of psychology, linguistics, anthropology, philosophy, computer science, and neurology. The paradigm is based upon computational thought (also known as the computer metaphor) which focuses upon the symbolic manipulation of 'mental models' of an external domain expressed in an internal 'well-formed language of thought analogous to a computer programme (Hunt, 1989).

This is an explanation of the phenomena, and it is only as good as an analogy goes. Certain authors have questioned whether a programme suffices as a model of thinking (Searle, 1980). The fields of information processing regards the programme and the system architecture associated with it. Knowledge representation plays a large part in this field, and attempts to examine the way the world is captured by mental models by mainly that which is conscious (Hunt 1989). In this spirit 'intelligent' programmes were given birth.

There have been several attempts to marry the principles and technology of artificial intelligence with computer assisted instruction. This has lead to the development of ITS's or intelligent tutoring systems (Sleeman & Brown, 1982; Wenger,
1987; Kearsley, 1987). The recent development of intelligent programmes for tutoring entreats a question that all technology used in education begs. If such devices are necessary and useful, what should their function be? How does one assess the efficiency and efficacy of such a system? How much teaching responsibility should the technology assume? How are you going to teach a subject or skill with the programme? One should have an idea of the level of student learning that has already taken place within that area of the material and what will be needed to be done at that point in time. There should be an appraisal of how much of the material is stand alone and how much is embedded in context. There is a need to appraise the helpfulness or interference of the use of devices, demonstrations, actual practices, or simulations. One of the main issues surrounding the use of computers in educational technology concerns the man/machine interface. This is often considered in terms of human performance factors centreing on questions of perception, attention, and memory, and their effect on control, workload, selection, and error (Wickens, 1984). The issues raised may be the foci of a formative evaluation (Weston & Cranton, 1986) of the technology to reveal the user/computer management dynamics of the learning process. Basing instruction on computers has taken into account task analyses of student learning in order to facilitate those tasks involved in the user-learning of a curriculum. The developers of intelligent computer assisted instruction, on the other hand, being more interested in artificial intelligence, tend to focus first on the 'knowledge representation' model in order to develop their knowledge bases. Kearsley (1987) proposed that one should implement both a system built up from principles of AI and task analyses. For example, Card, Moran & Newell (1983) propose the use of sets of goals, operators, and methods for achieving goals, and a set of selection rules for competing achievement goals.
THE FOCUS

One of the primary goals of software and hardware design is to allow the user to be in control of the power of the programme and the hardware in the shortest period of time and with the greatest ease. I wish to investigate how subjects react to, and behave or learn in a particular type of human/computer interface on a tutoring system — a microworld. Sleeman, (1987) describes this system as an environment which allows the student to 'explore' a closed world and yet provide (tutorial) support when it is required. A microworld is a computer environment which mirrors a small bounded slice of the real world (Papert, 1980). Typically, user actions in a microworld do not lead to intractable dead ends; all operations are reversible, thus reducing subject frustration and confusion. Rather than simply referring to "object" domains with "nice" representational properties, there should exist in a true microworld homomorphic mappings to structure which are representations of concrete action in the "real-world" (Groen, 1985). The system is modular in order that many different components can be accessed and used at any one time. This is important in a domain like medicine (Patel & Groen, 1986) where information comes from many sources (e.g.patient reports, laboratory results) and in different forms. The interface, in the research described in this thesis, is the XEROX desktop window environment. Winograd and Flores (1986) talk about the transparency of this sort of Macintosh-like window environment, where one does not have the extra task of programming, and which reduces the subject's need to have programming skills. These salient computer operations make operating a computer programme very much like other human experiences — one is conscious of the object of one's perception rather than the media through which one perceives it. Therefore, it behooves one who wishes to study reasoning (by operating a problem on a computer screen) to present the subject with a reasoning problem that requires the same sort of normal problem-solving as in the actual domain, while reducing the awareness of the medium of computer programming.
Therefore, in a microworld, subjects will not spend undue effort in the handling of the computer operating system.

COMMUNICATION

This study focuses on conversational discourse conventions in medical consultations and their effect, not mainly on subjects' understanding of the task they have to do, but on contextualizing the task they have to perform. That is the "connotative colouring" of the situation and the context in which the task takes place. Can subtle phrasing, even those which are for the most part unambiguous, affect reasoning? It especially notes how subjects respond to questions and to things that seem like questions or requests (Halliday & Hasan, 1976).

Every human being is active in a social dimension, in the causes that give rise to one's actions as well as to the purpose of the actions. This is true as well about conversation, especially the therapeutic interview, like the one between doctor and patient. There are sociological and socio-linguistic issues in medical practice surrounding the interactions and relationships between doctors and patients and between doctors and doctors (Bennett, 1976). For example, Fisher, (1983) showed that the asymmetrical relationship between doctors and patients allowed the doctors, not only to structure the interview, but also to propose treatments in an unbalanced manner. In this study, women patients showing signs of a certain cervical abnormality, who could control the topic of conversation, where able to avoid unnecessary hysterectomies for one of two other less invasive therapies.

The interactive relationship between the consulting expert's communication about his or her mental concepts and the knowledge engineer, who wants to operationalize the symbols in the eventual running programme are also important considerations. As well, in terms of teaching programmes the pedagog and the student may be a relationship to keep in mind, or one can restrict oneself to what kind of interactive surrogate the programmes interface can provide, in terms of reaching
certain educational objectives, by the application of certain methods and techniques. Paul Johnson (1983) points out the fact that people do not reliably express what they know. Katharina Morik (1989) claims that it is a leap of faith to believe that one has captured the expert's model in his or her head into a programme's knowledge base. The problem with knowledge acquisition is to elicit the knowledge from the expert using interview techniques, content and protocol analysis. These interactions require social and communicative skills which cannot be ascribed to computers today. All this is compelling since knowledge and its representation are important as well, and the focus of much current investigation. Researchers have proposed using structured interviews as a means to have some control over the type of information given by subjects (Easley (Jr.), 1974; Ericsson & Simon, 1980). This is to avoid introspectionism and capture reliable intuitions.

As this study focuses on aspects of the reasoning processes of medical professionals interacting with NEOMYCIN, an intelligent tutoring system (ITS), let us then take a brief look at the reasoning process of medical professionals.

THE KNOWLEDGE & THE SKILL

In reference to studying reasoning, problem-solving studies have provided evidence that there is no difference in the units of memory items which experts can hold in working memory versus novices. Experts as well as novices are restricted by the number of items on data in short term memory: $7 \pm 2$ items (Miller, 1956). However, Chase & Ericsson showed that experts have greater recall than novices. Thus, experts differ from novices with regards to the chunking of selective material (Chase & Simon, 1973). The difference is that the units in an expert's memory for his or her area of specialization are more sophisticated and compiled into interrelated units - i.e., chunks. It is this process of enhancing units of memory which is referred to as chunking. Chunking is a process of grouping interrelated information in memory into meaningful units as a result of iterative experience with stimuli (Anderson, 1985). The
selective inferencing which experts perform is more closely associated with viable problems (Chi, Glaser & Rees, 1982). They have acquired this skill via many thousands of hours of experience. However, in an unnaturally contrived, meaningless, or random situation, experts will revert to general problem-solving strategies similar to novices (Patel, Evans & Groen, 1989). Norman et al., (1979) upon examining clinical expertise, revealed that although experts did recall more typical clinical manifestations, than students, the more experienced physicians did not recall atypical findings differently from students. Experienced doctors need an extensive use of contextual inference in order to solve diagnostic problems, at least when no additional information regarding the present condition of the patient is available. This evaluatory condition forms sort of a constraint on their mental representation (Hobus, Schmidt, Boshuizen, & Patel, 1987). Different methodologies have indicated the differences between experts and novices in medicine. For example, Patel, Groen & Frederiksen,(1986) found that experts showed better selective encoding and greater use of prior knowledge when they read general medical texts.

Work in medical problem-solving often uses this expert-novice paradigm. This present study will, in part, observe some non-medical but conversational cues, which may affect medical reasoning.

The automatic reasoning of an expert is the culmination of many years of domain-specific experience which results in the development of a rich interconnected network of procedural and declarative knowledge—i.e., schemata (Anderson, 1985). Automatic tasks are triggered directly or indirectly by the attention process. Expert knowledge is highly compiled, thus permitting efficient, accurate problem-solving performance. However, this makes an expert's knowledge somewhat difficult to analyse (Anderson, 1983; Kuipers & Kassirer, 1984), as it is very quickly executed. Moreover, the expert has little metacognitive knowledge of the specific details of his or her reasoning. This makes the introspective reporting of their processing a
reconstruction – an explanation of the subjects' representation of their declarative knowledge. They can tell the psychologist what they “think” they have done, rather than what they “know” they have done. Human explanation is often based on idiosyncratic personal knowledge; moreover, it is willing to leave certain things unexplained. Morik (1989) points out that when one asks experts to talk about their own (inexplicable, unconscious) skill as well as their (more or less conscious and explicable) knowledge, one has to use natural language in the interchange. An expert may be an expert in the domain in question, but the explanations which he or she generates about self-behaviour, or even the field (if the expert is not a pedagog) may be “naive”. Thus the theory acquired will be likewise naive, because the person might not be an expert decipherer of self-behaviour. In addition, the way one asks questions in the interview will affect how the subject speaks about the domain. This will have an impact on how the knowledge-base is constructed. What an expert understands by a knowledge engineer's questions is another matter. Since the representation of knowledge is difficult to deal with, and beyond the scope of this work, the pragmatic, seat of the pants, interpretation of experts and subjects is best dealt with by noting their intuitions. These are same intuitions that would be used by them to explain or teach what they know to someone else (Macnamara, 1986).

There are many ways to characterize the development of expertise (Chi, Glaser & Farr, 1988). We shall look at two.

One way would be to watch its acquisition as it is being built up from the absolute beginner (Elshout, Jansweijer, & Wielinga, 1986), to the novice (Feltovich, 1981) then to the intermediate (Kaufman & Patel, 1987; Schmidt, Boshuizen, Hobus, & Coughlin, 1988), and through to the different levels of expertise (Elstein, Shulman & Sprafka, in press). This can be done longitudinally or cross-sectionally.

Another way to look at expertise, as used in this study, one may examine the behaviours of expertise when it fails – as it breaks down by either presenting subjects
with competing tasks which compels them to put more conscious effort into their problem-solving, or by giving them very difficult problems. This latter task presents the subjects with some elements of learning and explanation.

Patel & Groen (1989) suggest that explanation plays a part in expertise. Part of expertise is to be able to give consistent and coherent explanation. To cultivate this skill in reading, Palincsar & Brown (1984) have employed an intervention on poor readers which encourages them to question the main points of a text, to clarify and resolve difficulties in understanding, to summarize and to make predictions about what will come next. Their results showed success in improving comprehension. Perkins & Salomon (1989) have posited that, in the instance of non-trivial learning in complex areas, a significant means of achievement may be explanation. Focusing on cognitive skills, and in contrast to the somewhat instantaneous cognitions of expert automaticity, the present study examines what they refer to as the "high-road" to transfer of skill and knowledge. That is, cognitive activities which are spotlighted are the ones which depend upon learners' deliberate mindful abstraction of a principle. People sometimes abstract principles in advance, keeping them in mind in anticipation of appropriate opportunities for application, or, in a new situation reaching back to prior experiences and abstracting from them principles that might be relevant.

Research has made use of think-aloud protocols as an experimental method (Ericsson and Simon, 1984). This methodology has also been applied in medicine (Johnson et al, 1981). If one asks subjects to explain their own actions and to critique NEOMYCIN's reasoning behaviour, it may encourage them to verbalize these so-called "high-road" principles. It is difficult to say in this latter instance if what we are seeing is the slowed down process we have sought for or some other compensatory mechanism that becomes active when there is interference with the actual process. Nonetheless, such a methodology should reveal much about the nature of the necessary and sufficient
information needed to solve a specific case, as well as other ancillary information and the role it played. This present study asks that the subjects make explanations, critiques, and predictions of information.

In this experiment, the subjects are presented with a diversity of task demands in the complex NEOMYCIN problem-solving environment. This situation may be sufficient to impede the fluidity of the expert's automated reasoning, to make it more transparent and amenable to analysis. However, the transparent microworld should not interfere negatively with the subjects' problem-solving by creating extraneous computing tasks. Therefore what should remain is only the effect of the experimental manipulation. Moreover, in these experiments the subjects are asked to perform relatively few computer-operational tasks.

**WHAT IS NEOMYCIN?**

The design and use of an Intelligent Tutoring Systems (ITS) is one of the most controversial areas of computer-based education. This study has been run on such a system replete with a highly accessible microworld environment — NEOMYCIN (Clancey & Letsinger, 1984). NEOMYCIN is one of the best known and most highly developed of such rule-based systems (an ITS system containing an instructional environment for medical reasoning). It is a computer programme that models a physician's diagnostic reasoning within a limited area of medicine, namely, meningitis and diseases or conditions that can be confused with meningitis (Clancey, 1987b). GUIDON II (in this case GUIDON-DEBUG) is the case-based teaching end of this programme (Clancey, 1988). GUIDON is, in fact, a rule-based expert system in its own right (Wenger, 1987). It has a great flexibility to explain its reasoning as well as to evaluate students' performance (Clancey & Shortliffe, 1984). GUIDON can be used for the simulation and representation of expert behaviour. It is the part of NEOMYCIN which makes explications of reasoning the reasoning strategy of the programme. A very manoeuvrable ITS environment for computer-based medical consultation
(Clancey & Letsinger, 1984), NEOMYCIN is both sophisticated and pedagogical. The combination of a reconfigured expert medical problem-solving expert system, and the tutoring programme attached to the expert system, there is also HERACLES an expert system shell using heuristics to solve medical cases with domain, problem state, and strategy knowledge, and ODYSSEUS, a user strategy modelling programme which searches for the programme’s explanations of questions asked by a user.

Charniak & McDermott (1985) discuss the computer language LISP, in which NEOMYCIN is written. They indicate that the use of this language in artificial intelligence is because of its flexibility and user-changeable syntax, which is oriented towards the manipulation of symbols. When it is said that a programme provides an 'explanation', it is not the same as saying it has solved the matter conclusively. Logic, which implies bringing together what was scattered, seems to be by-passed when abduction in the sense of 'to carry off illegally', is used as an inferencing process which essentially can link together 'concepts' with no logical rationale. Abduction can be modelled by a computer programme and can generate 'explanations', which can permit false conclusions, as in the case of some medical diagnoses. Thus, the models or representations which a LISP-reproduced abduction produces, can reflect someone's first intuitions that are not necessarily true, unlike deduction, which is legal inference. People can know rules without knowing their application and thus can make inferences, predictions, and explanations based on untrue, albeit, operationally successful premises. In the case of deduction, on the other hand, with true axioms the inferences drawn will also be true.

GUIDON provides users with an opportunity to have a self-directed exploration of the system’s “expert” knowledge, and provides graphic depictions of diagnostic processes in ‘patient-specific’ models. The ability of NEOMYCIN to explain itself relies upon a complex rule-based knowledge base. GUIDON’s interface is similar to the critiquing and explanation tasks required from the subjects in this study. Like similar
systems, GUIDON uses both general and domain-specific heuristics with the capacity to model 'expertise', explicates its reasoning strategies, and allow self-directed exploration of its expert knowledge (Glaser, 1990).

These are systems such as SOAR which, when domain-specific procedures fail, apply general methods (Laird, Rosenbloom, & Newell, 1984). GUIDON teaches medical heuristics, plans, and procedures through examples, explanations, and evaluations of medical students' diagnostic hypotheses. It is neither a reference, nor a textbook programme. The application does not implement, in its rule-base, the diagnostic reasoning used in life-threatening emergency situations, but rather the reasoning used in a careful pedagogical explanation of clinical reasoning, analogous to an academic physician instructing students on how to eliminate other possible diseases. It combines an expert system and a tutor with a variety of interactive windows which can show medical findings, hypotheses, and so on. It is an interactive intelligent tutoring system built around an expert system with its rule base. It is capable of asking questions of students, answering their queries, modelling user schema, as well as allowing users to test their hypotheses. Intelligent learning environments (ILE's), on the other hand, are simulations built around trying to teach a specific item, much like trouble-shooting in electronics (Lawler & Yazdani, 1987; White and Frederiksen, 1986).

The essential aspect of the design of NEOMYCIN is its representation of the diagnostic procedures as abstract tasks that capture whatever structural effect the problem solver is trying to have on its evolving model of the problem (Clancey & Shortliffe, 1984). By running a medical "Consultation" on NEOMYCIN, a particular patient's data base can be accessed to see how the programme develops and to carry out a diagnosis on it. NEOMYCIN can "ask for" and "use" the information of a case history in a patient file. It can ask questions and make clinical hypotheses by means of which it can perform a differential diagnosis, that is, an ordered list of the most likely diagnostic possibilities for a particular medical case. The programme reasons
diagnostically using a directionally-oriented search strategy (from antecedents to consequences), termed "forward-chaining" that is intended to model the automated expert reasoning of human beings (also called "forward reasoning") (Greeno & Simon, 1986). The system's disease model and the heuristic procedure are intended to make NEOMYCIN's diagnostic procedures similar to those of expert physicians.

Forward reasoning is a search strategy characterized by moving from cues or data in the problem towards a solution (Larkin, McDermott, Simon & Simon, 1980). NEOMYCIN models this expert process. Forward reasoning has been observed in various experts: Lave et al., (1984); grocery shopping, Scribner, (1984); dairy working, Groen & Patel, (1986); medicine. Groen and Patel (1986) found that experienced physicians did not need to backtrack, unless their line of reasoning failed. Thousands of hours of experience have produced refined representations of goals and subgoals which can be exploited by experts to produce links of inferencing in the path of a solution state that would overload the relatively impoverished representations aiding the working memory of novices (Larkin et al., 1980). NEOMYCIN's diagnosing, therefore, should cause very little disturbance to the medical reasoning of competent medical professionals who observe it solving a case. Forward reasoning is disrupted by elements in a problem that produce uncertainty and necessitate back-tracking (Groen & Patel, in press). Although NEOMYCIN uses meta-rules which are outside of the domain of medicine, and certainty factors which are not transparent for physicians, these "mechanical" aspect of NEOMYCIN should not produce uncertainty in the subjects because these elements remain opaque.

The analysis in this thesis thus focuses on the considerations made during particular instances of uncertainty produced by a non-trivial medical case and computer-produced personage. These personages were subject-perceived or actual computer characters (in this case, "doctor" and "patient") who communicate with the user. The analysis will focus on certain pragmatic conversational elements of the
subjects' computer anthropomorphisation which may subtly affect their medical reasoning — as the subjects should have to infer "who" is giving them information: the data-base 'patient' or the computer-consultant 'physician'. Thus, the subjects comments about their reasoning processes will be partly the result of expressing cognitions surrounding these uncertainties. It should be noted that this does not refer necessarily to what the doctors may be confused about, but more broadly, to what they cannot know with certitude and as a consequence must infer.

**MYCIN**

MYCIN (Shortliffe, 1976), NEOMYCIN's predecessor, specialized in the diagnosis and treatment of certain bacterial infections. It had a differently configured knowledge base, and used "backward chaining" (Buchanan & Shortliffe, 1984). Backward chaining is somewhat analogous to "backward reasoning" (Greeno & Simon, 1986; Groen & Patel, 1988), which describes the several types of reasoning by hypothesis-testing which occur in novices (Larkin, McDermott, Simon, & Simon, 1980; Owen & Sweller, 1985), and in experts when they experience difficulty in a problem. The way MYCIN was programmed, overly compiled the underlying rules by which it made its decisions. The rules would have to be unravelled because they were what a student needed to learn (Clancey & Shortliffe, 1984). Adding the tutor GUIDON to the expert system MYCIN, as a pedagogical tool, was proven to be intractable because MYCIN's 'purpose' behind its reasoning could not be stated explicitly in a proper pedagogical manner (Clancey & Letsinger, 1984).

MYCIN's backward-chaining solves problems in a manner uncharacteristic to that in which expert physicians make diagnoses. NEOMYCIN's decompilation of the knowledge-base and its forward-chaining allow the rationale of the programme's behaviour to be made explicit.

In contrast MYCIN's reasoning procedure are: if deductive hypotheses states \( v \) then backward through evidence. It is easier to learn through the forward-chained
'explanations'. 'Explanation' as used here means to display rule justification which relates data/hypothesis associations, or hypotheses. The programme does not explicitly explain underlying causal processes such as pathophysiological mechanisms. The goal of the programme is to explain a diagnosis by indicating its relationship to clinical symptoms (Clancey, 1988).

One difference between MYCIN and NEOMYCIN exists at the level at which they can deal with language. Due to the hypotheses-directed backward-chaining of MYCIN, it produces focused questions which users can answer relatively unambiguously. Winston (1984) is of the opinion that a forward-chained programme may jump around producing user irritation. However, Clancey (1987a, p. 198) points out that while "At a certain level MYCIN is aphasic – able to perform, but unable to talk about what it knows." GUIDON and NEOMYCIN, despite data-directed forward processing, is a flexible communicator, responding to student hypotheses. It will be interesting to note how the language in NEOMYCIN affects the subjects.

NATURAL LANGUAGE

Millikan (1984) refers, grossly, to three orientations of studying the meaning of language:

1) Tarski, Frege and Davidson study the kinds of words, in terms of their effect upon the truth condition of the sentence (Frege deals with the semantic value, intentional content, and the propositional content).

2) Grice Schiffer, and Lewis try to describe the intent of the speech by means of declaratives and imperatives.

& 3) Austin and Searle examine conversational rules which govern the use of words and devices.

Studies of comprehension and problem-solving play a big part in educational psychology, and they should, also be involved in the study of nonstochastic, intentional decision-making. Language is an expression of cognition: a window to the
mind. It is an acquired intellectual habit. Language is a complex function involving several different faculties which accomplish specific and diverse acts; acoustic, phonetic, physiological, etc. Understanding sentences requires more than knowledge of phonology, grammar and a lexicon. The relationship between concepts and reality should be part of psychological investigations, when linguistic pragmatics are being studied. This is so because the focus is on how human beings think about, understand and use language.

Gottlob Frege (1949) differentiated between meaning (Sinn (sense)) and reference (Bedeutung). The latter construct refers to the real world; the thing signified — res significata. One cannot know a sign without knowing what is signified. Sinn doesn’t really deal with meaning, per se, but rather with a kind of meaning — modus significandi. That is that “sense” is a term in linguistics which signifies the same reality as the Bedeutung. This would be, for example, the difference between the word “dog” and a particular dog. Through this differentiation, Frege discriminates the concept from the object. If language is separated from its reference to concepts it loses all meaning.

Ferdinand de Saussure (1986) felt that in order for linguistics to be an empirical science, just as the material sciences were, language must be treated as an object. This perspective has led to great discoveries about the structure of language. De Saussure was aware that the whole is greater that the mere sum of its parts. He felt that the science of language (sign) was only one element in a larger science called semiology. By dealing with language as a faculty, without reference to the intellect or mind, and without regard to its origins or relation to the mind, de Saussure defined the object of modern linguistics as we know it. He did not feel that the object of linguistics was to study what he termed PAROLE — discourse, which is the concrete individual act produced each time a person communicates with a listener. The reason for this is that this interaction is of a multiple nature, and is contingent upon many things, about which a study of
language, *per se*, can have nothing to say. He referred to language as the *langue* — language, per se, can have nothing to say. He referred to language as the LA LANGUE — that is, the set of signs and the relations between them. This, and only this, he claimed, could be called English, French, or whatever. Thus, for him, language is a system for interrelations between signs, where the elements in the system cannot be defined in isolation, but only by means of their reciprocal interrelations. This would imply that the only value of linguistic symbols is pure relation.

The physical means of expression' was not how Saussure defined sign, as he was not consistent in his use of the term "sign". He began by saying that he would not be using the everyday definition of sign (a physical aspect of a thing which brings to mind some other thing) and brings both physical and mental attributes into signs, but later he uses the word in its everyday meaning. (Saussure, 1986).

There are other dichotomies in his theory, but here, for the sake of brevity, I will only focus on two others, for the sake of parsimony. Synchrony is a language at any given moment in time, and diachrony is language in evolution or change. He believed that to study language one had to look at it at one moment in history. The last polar concepts which we will discuss of his are that of signifier and the signified.

If one is to avoid a reductionism, in the model the SIGNIFIER is the word which immediately signifies the signified. The signified is the intelligible content of the world, otherwise known as the concept. Concepts are both signs and what is signified. The SIGN, signum, is the linguistic sign which is conventional and arbitrary. Linguistic signs are the union between the signifier signans (e.g., sound of a word) and the signified signatum (i.e. concept). The concept is an essential element of a linguistics of signs; its loss would denature language. The object is the reality. Therefore, the word refers to reality though concepts.

Many structuralists, including de Saussure claim that words do not have meaning except in contexts such as the deixis; "She is here."
This seems counter-intuitive. How then do interlocutors bring sensible form into consciousness, and make it their own? What are the conditions for utterance conventions when the utterance is appropriately made? This would be reasonable, if one typically wrote an isolated word such as "duck" and asked someone if it were a noun or a verb, but this is not the general purpose of words. It is true that one can understand the sense of a word, in relation to its actual context, and the context in which, and about which, it was uttered. One can say that words get their meaning from sentences or external situations since sentences do convey meaning, and listeners in new and unknown situations can communicate and understand, but words have a meaning of their own which they obviously contribute to the sentence. Context can be used to resolve ambiguity, but this does not tell us how it is used (Katz & Fodor, 1963). There are, no doubt, languages which can have one word for every sense, but English which is as natural as any language can be, is not like this. Language has to have a one to one correspondence between the mental and the physical for communication, as we understand it, to be possible (Hirtle, 1989). A statement can be independent of its method of verification.

Gustave Guillaume (1984) claimed that la langue is a prerequisite condition for la parole (Hewson, 1976). When we speak we first have the concepts, which we have learned and carry with us. They do not have to be re-learned every time they are used. This is in contract with Barsalou (1983, 1987) who claims we develop concepts as we go along. If this is so, then, what intellectual faculty do we do this with, and to what extent does context influence our thoughts? If we don't have concepts (internal representations of reality) how do we deal with phenomena from one moment to the next? Our experience has to be conceptualized in our minds, outside of language. If language has content, and it does, we talk about what we "know". Meaning, however, is not the same as knowing. Language is a mechanism which produces words and eventually sentences. The tongue must exist before an act of language can produce a
discourse. It follows that the prior conditions of all contextual meanings of a word is in tongue, and not among the contextual meanings in discourse (Hirtle, 1989). The meaning of a word exists before it is used in a proto-language (or a language of mind) (Eco et al., 1988). How then can thought, which man produces, and which does not have any of the qualities in common with his material body be considered to be material. Mental representations would require interpretation just as any other signs (Putnam, 1983). One may ask what sort of regularities are expressed in the semantics of representations for problem-solving in external domains in mentalese.

**MODELLING**

The approach of cognitive science sees meaning as intra-mental constructs. Our experience is not the meaning of a sentence, sentences have a meaning of their own. In analytical methods (such as the ones used in the cognitive sciences) one excludes, to a certain limited extent, linguistic meaning and experiential referent in order to create well-formed models of a discourse, story, or script. Analytical philosophy is characterized by styles of argument and investigation. It is thus possible for people of widely different perspectives to be practitioners of this sort of philosophy. It is a useful tool, but since language is a type of human behaviour, meaning and intentionality cannot be ignored completely. Thus, also, is the failure of reducing language to logic because the grammatical structure of philosophical propositions are not reflections of linguistic reality, and do not precede from linguistic conventions. Thus, although one can hold ontological realism at bay, while studying language as an object, one cannot exclude or deny meaning or the real world epistemologically in understanding language since it is linked with human knowledge and global understanding. That is, one cannot, at the same time, have a skepticism about one's knowledge of the real world and the nature of one's concepts 'capturing of reality', and simultaneously, feel that language is a reference to anything or has meaning. If so, it would mean that language would be no more a relationship of one person's ideations to another, in vacuo.
However a psychological analysis of human knowledge is not a philosophical analysis of reality (Gilson, 1937). Notwithstanding, the above what kind of ideas would we have in our minds if we had no experiential knowledge? Moreover, although language, as human faculty, may fall outside of the actual scope of linguistics, when that science studies language as a given, psychology does, conversely, investigate human intellect, making meaning possible, and cognitive acts, giving external expression through language.

Fodor (1981) feels that propositional attitudes play a very important role in psychology and in the philosophy of mind. A natural language set about propositional attitudes would include the beliefs, hopes, desires, wonderings, doubts, and so on of the cognitive agent. If language is the fruit of intellectual effort to give expression to one's knowledge of the world, then linguistic pragmatics is practising psychology. Whether one's knowledge be more or less complete or whether one's linguistic expressions are more or less suitable for the task of expressing knowledge, they are, nevertheless, based on external reality, with its own nature and properties. This is especially true when one is dealing with problems of intentionality and extensionality (Brentano, 1960; Geach, 1972; Marras, 1973). Reference to an individual should always be by way of essence, since individuals participate in diverse acts: being, the substantial form and many accidental forms. Any reference to accidental acts imply understanding that the act is sustained by the individual substance. One's logical knowledge of reality enables one to discern in what way a given proposition reflects reality. However, logic is not a branch of linguistics. Linguistics is the detailed study of language usage in an effort to see the nature of language. Language reflects the order of nature, not vice versa. Hence, real definitions (expressing the essence of an object) are not the same as verbal definitions, and thus, pseudo problems can occur in philosophy if one does not take into account the way that language functions (Wittgenstein, 1953)
If one were to reduce a linguistic expression of the real world to formal logical formulae, then one would ignore the genuine inference which links real facts for conditional links between propositions, and thus meaning would be prescinded.

The exercise of logical semantics ignores the relation between linguistic meaning and instrumental reference since it is not interested in mental processes. Truth conditional semantics ensure the objectivity and realism of their analysis by seeking meaning in the relationship language and the extra-mental reality. They have been criticized for being too abstract and not focussed enough on the actual act of speaking and understanding language. The linguistic approach posited by (Hirtle, 1990) seeks to observe the act of language, spoken or written, and its immediate surroundings. When we talk about words we have something in mind and then we say it. This is related also to our intuition of the tip-of-the-tongue message we intended to communicate— e.g., "I was going to say...". Thus, language provides a mental means of representing together with a physical means of expression — a word and sentence construction method. It is not static, but is an intuitive means for carrying out certain thought processes. Some claim that thought before it is verbalized is different from thought after it is expressed.

All one can try to think, concerning the not-yet-spoken, is conditioned by the impossibility of doing it without resorting to language. The only chance of observing thought in itself would be to catch a glimpse of it at the fleeting moment, when it descends into language, like the green ray the setting sun sends out just as it plunges into the sea. However, thought becoming language is already language, and besides, even if the experience of pure thought were possible, it would be necessary to make use of a language to communicate it. All one can do then is to work back up from spoken thought to that which is being spoken, doing one's best to discern the future thought of after-language in that which is being embodied. At best, this is difficult, because what what is poorly conceived cannot be expressed clearly, but there would not be any language if this could not be conceived at all. (Gilson, (1981)-
PRAGMATICS

Currently the word “meaning” is used vaguely. Haugland (1985) says that meaning can be thought of “as what is intended to be expressed”. It could also be the point or the purpose of the justification. It appears that part of the reasoning of people in reaction to language expressed in words is oriented to asking, “Why did he say that, or why did he put it that way?” Upon making a decision about this, each inference involves a supposition about the reasoning of the other person. People have understanding, in a pragmatic sense, revolving around beliefs, and intents, whose conclusions are not logically entailed (or at least not out of necessity) even in conjunction with common knowledge. Competence is the ability to make a rational decision in light of knowledge and perception. Therefore, this can also describe part of language competency.

TAKING FOR GRANTED WHAT WE KNOW

Understanding in instruction involves both comprehension and reasoning. Instruction incorporates comprehension and problem-solving. It involves, partly, the reasoned and deliberate considerations of an agent about how their actions lead to previously deliberated consequences. Both learning and problem-solving concern the acquisition, retention and use of new information, including both propositional knowledge, and procedural strategies. This is the realm of cognitive psychology. The components of a learning event, which are accessible for analysis, are the empirically observable task performance (through task analytical methods) and reasoning (cognitive skills). No one enters any situation, no matter how new, with a complete tabula rasa. There are biases which affect how one absorbs new information, using general knowledge and problem-solving skills. With respect with reasoning, the attainment of new knowledge and strategies alter the student’s beliefs, values and attitudes about the specific content and context of the domain under study.
One of the goals of instruction is to present material in such a way that is salient to discovery, in order to make the new configuration of concepts readily accessible to the student. Instructional methodologies work best if the student can be made to think. The student should be given information in such a way that it cues active participation in the learning process, rather than simply providing a list of facts to be received passively (Winne, 1989), and which can be regurgitated without assimilation. Intelligent tutoring systems, and intelligent learning environments habouring an environment for discovery, rule-base, or explanation-based learning, can create or simulate a 'real-world' environment with domain elements and relations in a microworld which can give the student domain-specific 'hands-on' experience in that situation (Papert, 1980). In the process of working in these environments, the student will develop metastrategies to manipulate items and relations effectively and efficiently within the domain. The eventual goal is that these skills will transfer into the actual domain as well.

At the early phase of complex learning there is much careful, effortful, strategic, and thoughtful behaviour which is accessible to metacognitive monitoring. There are also certain unanalyzed underlying assumptions which are made as the students start to develop a pattern or to use short cuts.

At the expert stage of a highly experienced, near automatic, problem-solver, many presumptive inferences are build into the reasoning. Although people seem unaware of what their reasoning is doing at this stage, it would not be far-fetched to propose that if such people were prompted for the mandatory use of certain cognitive processes, part of their cognitive energy might focus on the variables which with they employ to explain their reasoning. For instance, the cognitive strategy of explanation, which is usually seen in those with little prior knowledge (Tobias, 1989) could be used with more experienced subjects to see what they know about their rational assumptions.
As we learn, how much of our mental strategies become embedded untested beliefs? One way to expose these near ideologies is to challenge them. An indication that this has occurred would be surprise. Surprise is the normal reaction, to some person, place, event or thing which has not matched one's expectations.

Some of the most enigmatic assumptions made, however, revolve around the use of language. The assumptions combine both knowledge and meta-knowledge of a domain, as well as linguistic and meta-linguistic conventions.

H. P. Grice (1975) posits a set of conversational implicatures which he calls *Maxims of Conversations*. They could be seen as implicit considerations about order in the communication and the purpose of speech.

Grice's *Maxim of Quantity*:

A speaker/writer should make his or her contribution as informative (but not more informative) as necessary bringing forth new and unpredictable knowledge.

Grice's *Maxim of Quality*:

A speaker/writer should be truthful: not saying what one believes to be false or things of which one has inadequate information.

Grice's *Maxim of Relations*:

A speaker/writer should be relevant with reference to what is being talked about. Therefore, the communication is useful to attain some goal relevant to the topic.

Grice's *Maxim of Manner*:

A speaker/writer should be brief and orderly, avoiding obscurity and ambiguity and obeying proper form.

When a receiver feels that one of these Maxims has been violated, the assumption is that the communication has been altered on purpose, and thus the directly semantic interpretation of what was said, or read should take a second place to the meta-linguistic interpretation as to why the convention was altered. For instance, two people know what the time is, and they both know that the other knows. Further,
the other knows that they know. If one were to then say to the other, "Do you know that it is already nine o'clock?", it would not be felt that the communication was intended to indicate the time since this is an obvious fact to both parties. The communication might be interpreted to indicate an attitude about tardiness. Thus a reply from the recipient, would not be, "Yes.", but rather something akin to "Well, I am going as fast as I can." Therefore, breaking the convention of *quantity*, in this instance, implies or suggests that the producer means something distinct from what he or she said (Grice, 1975).

The Maxim of *Quantity* pertains to informativity. de Beaugrande & Dressler, (1981) use informativity in a slightly different way. For them, it is the extent to which the occurrences of presented texts are expected versus unexpected or known versus unknown or certain. Both refer to that which brings to the communication novel and unanticipated data. For tasks which require much reasoning, such as medical interviews, the focus of attention is on the information; signifying the content of new facts and their relations. Then assuming that the discourse is on the subject of *quality*, is truthful and relevant, it would follow that one would also assume that it is coherent. Thus, one can make a motivated search through the data to see what the occurrence of new information signifies in the problem-space.

Gazdar (1979) posits that receivers have presuppositions which can be cancelled by implications of a higher-order, which can be neutralized, in turn, by entailments. Thus contextual assumptions play a major role in interpretation after comprehension.

Therefore, beyond the simple decoding of text, there can be an element of interpretation focussing on what is not said. People tie together parts of the text. Bridging (Clark, 1975) describes the reasoning process whereby implications are drawn between phrases. Now reference cohesion, as Halliday & Hasan (1976) explain it, lies in semantic identity.
Reference cohesion can either be situational (exophora) or textual (endophora): anaphoric (phrases which refer to elements mentioned previous to the discourse and that co-specify with the phrase occurring previously in the discourse) or cataphoric (the opposite of anaphora; in which the connection is subsequent).

I make the basic assumption that some of a person's presumptions may be reflected in their verbal output. The textual manifestations of that presumption are substitution and ellipsis. Substitution, the replacement of one textual item with another, is a relation in the wording rather than in the meaning and ellipsis, the omission of an item, lies in grammatical relation in the text. In both substitution and ellipsis the entire clause which is substituted or omitted, is presupposed (Halliday & Hasan, 1976). Although they are used in the context of surface cohesion, they are used, in this study, to refer to the relation between parts of text, and their relation between meanings. Thus, the labels of types of clausal ellipses are used, in this study, to examine and indicate how subjects respond to questions, statements, and commands. Erotetic semantics (the semantics of questions) is the mechanical basis of this decision, since the analyses focus on responses to linguistic statements, and the cohesive relations between clauses.

One does not have the sufficient empirical proof to assume that all utterances are derived from implicitly complete sentences, however, at least a sentence can be written which would convey semantically the meaning attained, without having to make explicit all the obvious presuppositions, entailments, and implicatives. However this work treats these lexical decisions in a psychological manner, that is as an underlying interpretive choice which a subject has made reflected in the rejoinder uttered to the programme-presented text.

Traditionally, rhetoric dealt with the relation between the assessment and arrangement of text and its effect upon an audience. When certain usage inconsistencies occur when encountering a text, the use (in order to retain stability and
continuity of the ideas within the body of the text) must be logically cohesive and coherent in order for problem-solving to be viable. That is, a text must not only make sense in terms of its content, but in its ability to transfer the intended significance of the author to affect the reader (Miller & Glusksberg, 1988). In the experiments, in the present study, the subjects observe what may appear to be a dialogue, while they make it their task to perform a concurrent medical diagnosis with one of the 'characters'. With this in mind, two strategies of interlocutors in terms of text linguistics proposed by de Beaugrande & Dressler (1981) have to do with describing that which is being attended: self and the situation.

In the instance, when a receiver (the subject) does not share the outlook of a text producer (the programme):

**Strategy #2:** If someone else's monitoring does not match your own outlook, do not accept it.

You may (a) reject it
(b) question it, or
(c) ignore it
(d) replace it with your own monitoring (p. 174)

**Strategy #5:** Project your own desires and goals onto the other participant, except where there is evidence to the contrary (p. 176).

**Natural Language Environments**

NEOMYCIN does not have a natural language environment (NLE), which is compensated for by its ease of operativeness. Charniak & McDermott (1985) indicate that there is little agreement about the parsing of the English language. Ben Shneiderman (1987) points out also that, the problem with natural language interaction is not only its implementation on the computer, but also its desirability for large numbers of users doing a wide variety of tasks. How can one way of presenting language satisfy all users? In order for the interaction to be "natural", the programme has to be written on the assumption that the user is knowledgeable in that domain. One
wishes that the users can find clarification, if they need it, in a concise, clear, precise, pedagogically tenable, repeatable, consistent and expected way. The fact that ITS's focus on a narrow set of operations within highly specific domains, reduces the lexicon and thesaurus-like lists, and, restricts the linguistic interface to specifics. The more free and imaginative the text, the less able a system will be to handle it, and the less capable it is of giving clear direct feedback to users. Shneiderman (1987) states that computer users usually seek predictable responses and are discouraged if they must frequently engage in clarification dialogue. The production and difficulties in working with NLE's are quite formidable, and although NEOMYCIN is very user-friendly and transparent, there are consequences of syntactic ambiguity (sentence structure) and semantic ambiguity (i.e., inference related to entailment, logical implications and consequences, to name but a few). Theses factors may affect subject interpretation (Fodor, 1977; Shneiderman, 1987) and in fact how the physician may solve a case in this study. Although NEOMYCIN is not an NLE, the issue of language is important in terms of comprehension.

Let us take a brief look at systems which interface with users via natural language. Churchland, (1984) is of the opinion that within the domain of programming "intelligence", natural language production and comprehension are limited only because computer programmes do not have a knowledge of the world. Churchland oversimplifies the complexity of this issue because "knowledge" — the correspondence between what is in the mind and what the state of things are in actuality — material or otherwise, is a very knotty metaphysical nut to crack. He states that for many reasons it is desirable to converse with computers in ordinary English, and that natural language manipulation is a central element of human intelligence, in any case, worthy of simulation in its own right. It may indeed be worthy, but it is not trivial. It is one thing to manipulate symbols syntactically (well-formedness), it is entirely another thing to perform semantics (meaning).
Some of those reasons are pointed out in Wenger's (1987) discussion of SOPHIE. This is a programme by John Seely Brown and Richard R. Burton (Brown, Burton & Bell, 1974) that uses a parser as a language interface on a system which allows users to perform trouble-shooting on configurations of an electronic circuit board. At least one reason is the fact that the more powerful, cooperative and realistic a natural language interface is, the more domain-specific predictability and knowledge is expected from the user. Hence the scope of the linguistic interface has to be limited and restrictive, as the system has a limited ability to explain real-world causes, purposes, intentions, scripts, plans or goals; which natural interlocutors expect. Churchland (1984) points to Joseph Weizenbaum's ELIZA (a mock up of pad responses of a non-directive psychotherapist to a patient who types in responses) and Terry Winograd's programme SHRDLU (a simulated limited domain environment of blocks) as examples for the potentiality of simulating natural language (Winograd, 1972). His two examples, however, exemplify the hurdles. ELIZA denotes a system that "works" (passes the Turing test) in a very limited context. If the user varies slightly from the constraints, the system fails. Hence, it will give highly cooperative subjects expected and familiar sentence strings related enough to the small context that it fulfils the subjects' expectations (Dehn & Shank, 1982). For ELIZA, the interaction falls apart if the user doesn't play the role of a patient. It has no "intelligent" features and no knowledge base. SHRDLU, in contrast to ELIZA, operates 'objects' in a finite domain in order that the operations can be trivially monitored by the programme (Winograd & Flores, 1986; Suchman, 1987). Trivial does not imply that its well-defined operations cannot be extensive enough to carry out many procedures. However, the domain in this thesis is ill-defined and the structures which pertains to its subject matter are accessible only through elaborate heuristic (rule of thumb) formulas.

Even though great advances are being made in natural language processors, the ramifications of semantic and pragmatic ambiguity (inferences and interpretation of
linguistic expression as a function of non-linguistic content) must be heeded (Belnap & Steel 1976, Shneiderman, 1987). Semantic ambiguity refers to ambiguity about the meaning of a lexical string. The sentence "One very important thing which you must be careful never to forget is that you can never put too many anchovies in a Caesar's salad," either is intended to mean that one should be sparse with the use of anchovies or that one can be as liberal as one likes in the use of anchovies in this recipe; in fact, the second interpretation even implies a warning about using anchovies too sparingly. As in the game-theoretical semantics of Hintikka (1976) the concerns of pragmatics are dealt with by exploring the attitudes of producers (termed 'intentionality'). Intentionality refers to the mind's ability to refer to things beyond itself and to mental objects (Angeles, 1981). Generally, however, when the term intentional is used in this paper it refers to purposeful and directed behaviour; to having intelligence and will. De Beaugrande and Dressler, 1981, indicate this usage with respect to understanding the purpose behind texts. What they refer to as intentionality, concerns the texts producer's attitude that the set of occurrences with a text should constitute a cohesive and coherent text instrumental in fulfilling the producers goal e.g., to distribute knowledge or to attain a goal specified in a plan. This is whether or not the producer is the text presenter, which would not be the case in an ITS. De Beaugrande and Dressler, 1981, round out this idea by describing the term acceptability. Acceptability concerns the texts receiver's attitude that the set of occurrences should constitute a cohesive and coherent text having some use or relevance for the receiver, e.g., to acquire knowledge or provide cooperation in a plan.

Dennett's 'intentional stance' (1987) involves the use of ordinary mental concepts: belief, desire, hope, expectation, imagining, and the like for explaining behaviour. Thus, in the context of this thesis, it is used to describe the purpose behind a text in terms of propositional attitudes about the text producer — whether that be the writer, or a computer character. It does not refer to intentional species. Putnam, (1988)
points out that we cannot individuate concepts and belief without reference to the environment, since meanings are not "in the head". Propositional attitudes: such things as believing that snow is white and feeling certain that the cat is on the mat are not "states" of the human brain and nervous system considered in isolation from the social and nonhuman environment. This concern is not all that crucial with respect to the subjects in this thesis, since the subjects interpret the text in light of their experience with people and with medicine, and the circumstance they are in. Therefore, trying to determine what the text producer might have meant is just one more consideration among many environmental considerations. Pragmatics also deal with receivers — 'acceptability', and with the communicative settings — 'situationality' (as in model-theoretic semantics) (Barwise & Perry, 1983) (de Beaugrande and Dressler, 1981). Situations in model-theoretic semantics need not be consistent or maximal as they are in possible worlds (Kripke, 1980). Possible worlds are useful for looking at problems of uncertainty dealing with modal logic. With respect to knowledge representation, in the context of problem-solving it is a way of thinking of other ways this world of ours might be or might have been (Barwise & Perry, 1983). Although the model-theoretic approach brings in context, lack of consistency would make interpretation difficult as model-theoretics. It can account for entailment but fails to have realist links, since it is not dependent on how the world is configured, but only on the logic of the linguistic world (Barwise & Etchemond, 1989). The issues of presuppositions, opacity, implications, implicatures and the performatives in speech act theory (Austin, 1962; Searle 1969; Winograd and Flores, 1986) are contextual issues which affect pragmatic inference. Opacity is a problem of co-reference, implicatures are inferences base both on the content of what is being said and some assumptions about the cooperative nature of ordinary verbal interaction, all of these are actually cases of language usage, meaning and reference affecting language structure (Halliday & Hasan, 1976; Macnamara 1982, 1986; Barwise & Perry, 1983, Levinson, 1983).
Pragmatic ambiguity refers to presuppositions in terms of the discourse worlds about what is being referred to (Hintikka, 1976). That is, an indecision about what an utterance implicates. Consider the sentence "Plato is on the top shelf". The reference is certainly not to the late philosopher. Persons establish links between objects of a different nature for psychological, cultural, or locally pragmatic reasons. The established links allow reference to one object in terms of another (Fauconnier, 1985). Therefore, a remark like "That's a nice horse you have Bill. It would be a shame if anything were to happen to it," can be taken as an idle comment, a warning, or a veiled threat. What then can be said about what users say while they interact with computer programmes – about subjects' response to linguistic pragmatics conversational convention which connotes things which are not specifically denoted (Evans, 1985; Gazdar, 1979)? Will these things affect a subjects understanding or reaction to the computer information? There is no reason to believe that in complex computing environments this may not occur. Winograd and Flores (1986) make the point that as computers do not understand natural language. That is, they do not have semantics (meaning in its true sense) only symbol manipulation (Searle 1980). It is the subjects who are using the structure of their language to interact with that system and both the builders of the system and subjects may be unaware of what kinds of implicit assumptions have been imbued in that system (Wenger, 1987).

A great part of basic comprehension has to do with the interpretation of sentences (Macnamara, 1986). In real-life information and task domains, the interactive components between the recall of chunks (units of mental representations) of important information and comprehension are important. Moreover, communication between interlocutors are comprised of more than an exchange of semantic information. There also exists elements of intentions, time, and expectations. More specifically, subjects may read intentionality into a programme and alter their actions towards it in light of this. The feeling that the programme has a
will and intends certain actions should be evident in the subjects verbal reports. Macnamara, Govitrikar and Doan (1986) defined actions as events that can be truly explained in terms of the beliefs and desires of the agent that produces that event. A behaviour, on the other hand, is "the publicly observable activity of muscles or glands of external secretion" (Hebb, 1958 in Macnamara et al, 1986). Actions are affected by mental factors. Subjects need not be consciously aware of all the beliefs and desires that motivate their actions. Macnamara et al. (1986) stated that

"... causal laws to cover actions could not be discovered. We are not surprised, then, to find those areas of psychology that concern themselves with the explanation of action make no appeal to such laws. Instead they appeal to systems of beliefs and desires, whether conscious or unconscious. This is to explain actions in the only way they can be explained, by making explicit the motives that make the action intelligible. " (pp. 37-38).

If philosophers of mind explicate actions in terms of motives, it is also expected that the subjects in anthropomorphizing the programme may attribute intent to it the way people often accredit intellectual intent to animals. For instance, in the domain of tort law there is a concept foreseeable risk. Within this concept people are seen not merely to be responsible for their intent (such as in the case of recklessness), but moreover inasmuch as they are aware of the consequences of their actions. This includes conversation conventions. Intention can only be had by a person. One does not sue a dog, one sues its owner, and not because most dogs don't have any money. Colloquially, the dog "intended" to bite you, or it would not have lurched forth and sunk its teeth into you, but most would not believe that the dog had a plan, or that it had it out for you.

Looking to language as a window of the mind, some psycholinguists, who are interested in the development of language and in its pathology, offer some insights. Bloom and Lahey (1978) talk about three aspects of language: its form – the
orthographic, phonological, morphological, and syntactic aspects of language; the mental content of language — what people talk about and what they understand of what other people say, and its use — the mathetic and pragmatic functions of playing word games or making requests; and the linguistic, and nonlinguistic context which we react to. Respectively, in medicine, form, content, and use may correspond to the syntactic/grammar of the diagnostic string, the semantic medical domain of focus and the explicit denotation of the problem-solving procedure, and pragmatic issues (e.g., pertaining to the medical discourse). The investigation of this current thesis is oriented towards a scientific examination of subject behaviour and action describing the psychological actions of intentions, interpretations, and expectations from a practical problem-solving user-interface point of view.

**SEEING THINGS**

How does one go about analyzing discourse text and studying the intuition of subjects who are actively engaged in a mental task? In this study, the textual evaluation is only semi-analytical in nature. The actual excerpts of the transcripts are used to make points about the subjects' reactions to the interface. This, along with background on the basis of the physicians medical reasoning, allows the situation and the comment made by the subjects to speak for themselves. Any dispute about their interpretation would lie on that intuitive level. To fully understand language, one needs to have the metaphysics, the logic, and an understanding of language conventions in light of situations: from reality to thought and from thought to logic (Lena, 1959), which I certainly do not intend to address in the embryonic work.

One the other hand, with respect to the the pictorial analysis, it is often said that a picture is worth a thousand words. Well, in this case, the pictorial analysis is used to reduce the amount of elongated transcriptual reading. The diagrams focus on the actual medical reasoning, and the considerations which revolve around it— which is the focus
of this study. The conceptions are not decompositionalized, they are taken them as they would be understood by the subject or another physician.

CONSOLIDATION

The principal aim of this chapter has been to raise the issue that language use can reveal much about the way subjects reason in complex domain. In particular, the focus was on cognitive processes pertaining to the extraction of meaning which generally happen so quickly, and automatically, that they are taken for granted.

In review, generally, ITS's are designed to interact with the user without a need for outside intervention from any other individual. In order for an ITS to get through a tutoring session, it must be able to model the user, and orient itself in response to his or her activities within that system. In this study, it is important that the programme should interact with the subjects, sufficiently, so as to avoid having the experimenter continuously interrupt the subjects' progress by introducing extra verbiage. This is one reason why an ITS was chosen as a means to study human/computer interface issues related to interpretation, and understanding, as the subjects involved themselves in a task.

Educational psychology lies at a crossroad between practice and research. It pertains to the art of allowing individuals to achieve mastery in some domain or domains, and is involved, at the same time, in the study of what elements of our nature we can come to know which may facilitate this learning process. Operating in a microworld environment may serve as a tool for both of these purposes: at once, being a 'place' where learning may occur, as well as serving as a medium for research on reasoning.

The experiments in this thesis involve situations which a subject can become engrossed in, thus minimizing external distractions (e.g., experimental artifacts). This is akin to the ethnographic approach of placing the subjects in a situation where their awareness of being monitored is inconsequential. They therefore, act naturally. One of
the familiar elements of the interface, which the subjects will be involved with, is that it is medical in nature, which is the professional domain of the subjects. The experiments also involve situations of which the subjects are familiar, such as: talking with a patient, or consulting with another physician, or explaining a case to a third person. It appears likely, that as far as the similarity between solving a medical case on a computer goes, versus doing so in a natural setting, it will be to this extent that the subjects will bring the same cognitive considerations to both. The subjects will obviously lack a live patient, whom they can see and touch, and things are not happening in real time. Nevertheless, the only 'way into' this 'microworld' is through the text. It is a bit like a text adventure, in which the imagination serves to fill in what is not provided by the text. What the subjects bring to their imagination will have been influenced by their past experiences. The semantic sense of a lexical string can be affected by its syntax, but it is not syntax alone which adjudicates user-interpretation in situations. It is, rather, the things which the user will find typical which will drive their expectations. This, in turn, will affect their interpretation of the programme's text, beyond its literal meaning. The conversational conventions which people use, will have implications on how the text may be understood. Semiotic interpretation may lead to metasemantic inferences, some of which may be captured by pragmatic analyses of the subjects' think aloud protocols.

GUIDON serves as a communicator between the system and the subject to reveal information about its knowledge base to the user. As well as use GUIDON uses that knowledge base to carry out its own diagnoses. It is important for us to determine what kinds of unanalyzed expectations people may have, which affect their decision making – the kind of notions which precede the interpretation of evidence which one encounters.

The degree of order that a system, such as NEOMYCIN, can have on the presentation of verbal data is useful so as to control certain elements of the
experiments while leaving others free. Because the subjects, in this study, are given a definite tasks of problem-solving in a medical case, the direction and purpose of their activity on the computer is somewhat regulated. Moreover, their efforts to make sense of the text to acquire the whole clinical picture will be a problem-solving task in itself, which requires that they make inferences. Inferencing differs between experts & novices; with experts having more command of relevant inferencing in their domain-specific area of specialization.

The medical case which the subjects will encounter will be presented as a sequence of questions and answer. With this type of presentation, even an extremely quick expert would be likely to provide much verbal commentary not only about the case, but about the diagnoses performed by the programme. Requesting that the subjects verbalize their reasoning, while working with NEOMYCIN, should encourage them to make explicit their inferences which pertain to entailments, implications, implicatures, and presupposition which come about in their interpretation of the text and the situation.

Although NEOMYCIN does not have a natural language processor, it communicates with users in English sentences. The circumstances in which it uses language, as well as, the relevance of the text to the situation, is expected to be assessed by the subjects in the same way as they assess natural discourse. The subjects have to arrive at the meaning of the statements in terms of what they know about the patient and the situation, in particular, and what they know about medicine, in general.

What is NEOMYCIN doing in its operations to support this with respect to the anthropomorphisation of intelligence and volition into the interface, on the part of the physicians? There are two things. First, NEOMYCIN is asking diagnostic questions, the reasoning behind which, can be analyzed in context. Second, the subjects are accumulating additional information about the patient by virtue of the programme's response to the programme's own questions. In addition, NEOMYCIN's capacity to
explain itself, makes the interface highly interactive. The programme also serves as an environment where the subjects can perform certain operations; such as ask the programme for explanations.
CHAPTER III

This chapter describes the purpose of the study, followed by questions to be addressed in the study and the rationale. A series of hypotheses are generated at the end of the chapter.

PURPOSE OF THE STUDY

This study is designed to observe physicians' medical reasoning as they interact with a user-friendly computerised intelligent tutoring system (ITS). NEOMYCIN functioned as a consultant for the complex ill-defined problem-solving area of medicine. One reason for using this particular type of interactive programme environment is to see if it can be used as a way to elicit rich verbal think aloud protocols from subjects, which include the subjects' linguistic pragmatic considerations. The man/machine interface consists of an interaction between a doctor and a patient, which the subjects observe. It is when the subjects begin to address the characters in the interface. The subjects are anthropomorphising the interface, especially if they speak about the individual programme characters differently, or address characters which are not immediately present in the interface. Regardless of whether or not characters in the computer interface are personified, the text which is read by the subjects is open to interpretation beyond its denotative meaning. As a result subjects may read certain connotations into the programme's textual interchange. Subtle formal structures which may alter the function of wording (syntactic phrasing dynamics) may influence which interpretations are made. In addition, the operation which a subject performs to operate the programme consist of a series of subject actions and computer reactions, followed by subject reactions to the programme. This exchange itself, which is the experimental task, is open, as well, to subject interpretation; with the possibility of not clearly be understood by two individuals in the same way. The goal of the study is to determine how computer
anthropomorphisation and the corollary syntactic phrasing dynamics affect the pragmatic and semantic comprehension and the clarity of the task of a man/machine interaction in a microworld environment.

The operative task for the subjects is to survey NEOMYCIN (which is in the process of making a differential diagnosis of a clinical case) while they simultaneously make a parallel medical consultation of the same case. The domain content of this study is neural and cerebrovascular pathology and the intellectual task is clinical reasoning. The programme models two physicians consulting on a medical case. One 'physician' asks for data, and the other replies. The system is designed to create an interactive environment which resembles what a physician might encounter when consulting with an unhurried specialist regarding a clinical case. The subjects' procedural behaviour and verbal critiques in response to this environment should indicate how they use their knowledge of medicine, as well as their considerations of conversational conventions to make inferences.

Their verbal behaviour will be evaluated under the rubrics of 5 types of cognitive operations— reasoning: which are the subjects' clinical reasoning utterances; explanation: which are their spoken justifications for their actions; prediction: which is the response to the experimental instruction to forecast patient findings, and the direction of the programme's clinical inquiry; making a critique: which is the ongoing commentary the subjects make pertaining to the experimental trial, and anthropomorphisation: which are instances when subjects' verbal responses mark an attitude, as though the subjects were interacting with persons, rather than with a machine.

The first 4 operations could be stated as statements:

Talking about what I am doing: Reasoning.
Talking about why I am doing it: Explanation
Talking about what is foreseeable (as a consequence of what I think to be true): Prediction.

47
...and talking about my attitude about what is being done: Critiquing.

QUESTIONS TO BE INVESTIGATED IN THIS STUDY

1. Are the considerations of the physicians' subjective interpretation in linguistic exchange between themselves and the machine substantial enough so that the subjects anthropomorphise the interface?

2. Are there parallels between doctor/patient interaction and this man/machine interaction in this domain?

3. In an interactive environment, do professionals with varied experience show different types of clinical reasoning?

4. In the interaction do the cognitive tasks of explanation, prediction, and critiquing play a critical role in diagnostic reasoning?

FOUNDATIONAL ISSUES

There is an ever increasing reliance on the deployment and usage of "intelligent" technology. This technology mediates between ourselves and the operations which we wish to perform. We must look, not only at what these machine can do, but how they are being used by us. How do we interact with computers? Do we have the same expectations of "intelligent" software as we do with human professionals? As Winograd and Flores (1989) stated we should take advantage of the machine not for its ability to appear human, but for the fact that it is not human and possesses nuances that are useful to us and lacks other aspects which one may find a nuisance when trying to do similar interactions among human beings, and vice versa. The point is that the many things that make an interface "friendly" beg the question as to why? The answer is that the machine does what we expect it to do. How, and in what manner then do we rely upon computers to do "intelligent" things for us? What is the nature of our expectations? They probably come from our everyday interpersonal communication, and as sophisticated machinery allows us to keep these expectations, we should know what they are. This is the realm of social psychology. Man/machine
interface misinterpretation can be due to attempting to “reason” with a machine. Thus, the issue in this study is how the user reads meaning into things. For a complex interactive interface, one must recognize the possibility of signalling messages that transcend the intent of the programmer being interpreted by the user. This is because the programmer is working with the same unanalyzed socio-psycholinguistic expectations and assumptions.

By having experimental subjects critique a computerized model of an expert physician, the subjects are anticipated to disclose more of their knowledge about clinical reasoning; moreover, the influence of pragmatic assumptions on semantic inferences may be made ostensible.

In this study, NEOMYCIN presents information for a case that should allow any physician to make a medical diagnosis of the case. Notwithstanding, the case is not an easy one, and not in the immediate area of specialization (neurology) of any of the subjects. Thus, the subjects may be described as subexperts—expert physicians reasoning about cases that fall outside their domain of expertise (Patel, Arocha, & Groen, 1986, 1987).

Subjects were selected in this NEOMYCIN experiment such that there be some learning or exploration on the part of the subjects via the experimental task. This was carried out in order that questions physicians had would be more likely addressed towards the clinical case rather than about the computer itself. It was felt that expert physicians might feel challenged by the machine and would not perform normally; assuming the experiment to be a test of their clinical competence. More importantly, however, results have shown that subexperts tend to use mixtures of both forward and backward strategies. The use of backward strategies seem to occur whenever a physician entertains several hypotheses at once or sequentially and then checks their legitimacy against the information provided by the patient (Patel, Arocha, Groen, 1986, 1987). Although it would be difficult to determine the directionality of the
subjects' reasoning, this type of explicative and critical reasoning should provide us with the type of think aloud protocols that we are looking for, with respect to intentions, plans, scripts, goals and assumptions which we are looking for (Suchman, 1987).

We expect to find a number of things from the think aloud protocols collected in this study one may find a number of things: The quality of the on-line think aloud verbal protocols, to see if presenting text in this manner induces physicians to think-aloud about matters which include aspects of the human-computer interface, and the anthropomorphisation of the programme. Specifically, the subjects may not always perceive two physicians having a conversation, but sometimes a patient volunteering information, or as feared, a test environment produced by a psychologist/experimenter where they are being giving a "tricky case" with an experimental hidden agenda.

In order to attain certain contextual information from the participating physicians, these subjects are probed for it. Before the experiment, the physicians will be told, in part, what they were being probed for. There is always the danger that this might change the way they will reason. There is a trade-off here. If they were not probed, they might spend much of their time keeping their ideas silently to themselves. Nevertheless, there is a conflict between necessity of pre-experimental probes (instructions) and their effects (good or bad) on reasoning. Since the computer system is designed to model at each step what an expert physician may have done, it is thought that it may give some insight into how real physicians go about solving a problem as they deal with a intelligent tutoring problem, which is seemingly making considerations similar to his or her own.
HYPOTHESES FOR THIS STUDY

The hypotheses about the outcome of this research are as follows:

1. The physicians may personify certain characters, such as patients, other physicians, or evaluators of their competency. This is due to the arguments posited above that the subjects will treat "intelligent" things as though they had intent. That is, the perception of the behaviour of the machine rests upon a desire to produce an effect which has a goal, in this case to intelligibly solve the case and to be communicatively cooperative with the user.

   - If the subjects personify characters in the interface, they will not give the same weight to evidence which they perceive as coming from a patient to that one they perceive as coming from another physician. This is assumed, as one may purport that belief rest upon one's internal representation of the world, and the representation of medicine differs between patients and medical physicians.

2. There will be similarities to doctor-doctor and doctor/patient discourse in the man/machine discourse. The physicians may elicit analogous situational behaviour. That is they may act, to a degree, as though they were in an actual clinical setting with actual parties.

   - However, the sentential formulation of the programmes' reply, may not always fulfil the expectations of the subjects' inquiry and produce objections from the subjects. Much in the same way a request for a cup of coffee does not constitute a request for crystals floating on the surface of cold water, the nature of the subjects' objections may reveal the kinds of presuppositions which they employed to interpret the interaction.

   - When engaging in GUIDON's explanation facility, the exchange can take the form of a genuine human request/computer explanation dialogue. However, when a subject requests an explanation, he or she is constrained by the limitations of the interface not being a natural language processor. That is to say that, in this study, the
subject can only make an unspecified request for an explanation, by selecting explain. The subject cannot request the type or the direction of the explanation. The subjects would have to rely on the programme answering a question only around the rules used to explain its current finding focus. GUIDON's 'explanation' facility will transmit the critiquing into an interactive task and so allow critiquing to build on itself. When the subjects makes an inquiry for an explanation, it is like asking the programme why it is taking a particular route. The programme's reply are its explanations.

3. Explanation, prediction, and critiquing will allow one to examine the subjects' problem representations, as they focus the subjects' verbal protocols to these aspects of their clinical considerations.

- In addition to the medical reasoning observed on the subjects' think aloud protocols, specific evidence of the subjects' clinical reasoning will come from their explanations and their critiquing of the clinical reasoning of NEOMYCIN. They are probed for their predictions as a way to determine their representations of the problem at a given moment of time. (It is possible that NEOMYCIN being an "expert" may lead the subjects – themselves not being specialists in neurology and infectious disease – through a pedagogical tour, without the subjects generating any diagnostic hypotheses of their own). This would imply a relative lack of critiquing. However, in terms of active versus passive participation, critiquing is a criterion to denote active participation, and it is hypothesized that the subjects will participate actively.

- Since NEOMYCIN is not in "emergency mode" physicians may feel it is being too academic in its orientation and not pursuing at every moment the necessary investigation for the saving of the patient's life. If subjects perceive a state of urgency to solve the case before the case gets worse in terms of immediate treatment, they will have conflicting goals with NEOMYCIN.

- The subjects may perceive NEOMYCIN as a test environment produced by the psychologist/experimenter where they are being given a "trick case" with an
experimental hidden agenda: an electronic Skinner box, of sorts to monitor technophobia in the medical professional, perhaps.

The following chapter illustrates the nature and elements of the particular problem-solving task which the subjects' encounter. It outlines the medical case and the reasoning decisions that are entailed when diagnosing it.
CHAPTER IV

BACKGROUND: THE MEDICAL CASE

This chapter describes the elements related to the domain-specific reasoning required in the experimental trial — meningitis and diseases and disorder which may be confused with it. It begins by describing what is a problem space and then goes on to describe the normal system and pathological conditions which may arise. It ends with a diagnosis of the case provided by a set of consulting expert physicians: an infectious disease specialist, two neurologists, and an epidemiologist, who was a nurse; all of whom are specialists with respect to the domain of the case.

DIFFERENTIAL DIAGNOSES

The physicians cognitive representation (domain knowledge and its internal organization) of the problem is established during operating in the problem space. This is propelled by general or domain-specific knowledge. This chapter identifies parameters of the potential problem-spaces for the case: Debra — the patient case situation which the physicians encounter in this experiment. Within the system of information processing (Newell & Simon, 1972) problem-solving procedures are characterized by search in a problem-space (Newell, 1980). A problem-space describes the cognitive framework made up of a array of knowledge states and operators to transform these states. Problem-spaces consist of a set of initial states — in this instance, Debra's presenting illness. It also comprises sets of goals, states and a set of path constraints. The subjects, who are doctors, have to make diagnoses descrying knowledge from finite information about an arrangement of disorders bearing protean symptoms. Ill-structured problems, such as those which exist in law and medicine, require knowledge from several different sources, and are so, because the solution requires coordinated work in several disparate problem spaces. A problem can be ill-
defined, as well if the goals which are formulated are underdetermined; i.e., several alternatives may be specified (Simon 1973).

Good physicians keep their hypotheses broad and vague as well as multi-faceted and will change their reasoning path, if they feel they are getting nowhere with a line of inquiry (Barrows & Bennet, 1972). The term diagnosis refers to either an operative procedure effected to make considerations about symptoms and signs in order to determine underlying diseases, or to the end judgment achieved by that process. This study concerns itself with the former consideration. NEOMYCIN is, in its own right, a diagnosing instrument the domain of which is meningitis and diseases which can be confused with meningitis.

THE TASK

Following Ginsburg (1980), it is from a body of possible hypotheses that a single diagnosis must be deduced from 1) history-taking – the disclosure obtained from the patient's description of her life, her physical well-being, her mishaps, medical interventions, patient identification, the reason the patient sought medical help and illnesses. It also looks at personal and lifestyle habits. It is interesting to note that in the first experiment unlike a real clinical interview the subject does not regulate the verbalization between himself and the patient; as is characteristic of this task, but he or she does control the inquiry in the second experiment. 2) Further information is obtained through physical examination. Error in a medical diagnosis may occur as a result of improper history-taking more than missing something on physical examination or laboratory tests. 3) Data obtained from routine laboratory examinations and special tests help to resolve the patient's clinical problem which requires that the cause of the problem be defined as clearly as possible in order to select appropriate means of therapy (Ginsburg, 1980). However, some therapy often has to be implemented before laboratory results are available. The first step in the clinical reasoning process is to decide which of the patient's preliminary symptoms and signs
are worth pursuing for possible therapeutic requirements for their diagnostic and prognostic potential. It is very important not to be side-tracked into a premature analysis of the first symptoms forthcoming, but rather to obtain an overview of a patient's symptoms and problems, listing them, if necessary, as the patient speaks, and selecting those that one wishes to evaluate (Ginsburg, 1980). Some of the factors that contribute to the decision about selection of a problem include: the nature of a finding (some findings can never be overlooked and demand explanations); the severity of a finding, the course and duration of the finding, the clinical background and the extent to which the condition is bothering the patient. Once a problem has been formulated, one needs to know its common causes to determine what is its cause, and be aware of the more probable causes in a particular patient, in view of that patient's known medical background. One needs to assess its consequences: know their clinical manifestations, and be able to detect them.

THE PROBLEM

To identify the medical problem-solving task which the physicians face in this experiment, it may be useful to discuss the types of meningitis and the different mechanisms: infectious, vascular, immunological, toxic, neoplastic and psychogenic that can be confused with it. The patient, "Debra", whom the 'physician' "sees" in this experiment, presents with symptoms which indicate a meningitis. An accurate diagnosis of meningitis is important because untreated its natural course is dismal; however, it is very amenable to treatment. Treatments make a difference in the natural history of the disease; moreover, morbidity and mortality will be greatly reduced by the appropriate diagnosis and treatment of the disorder.

THE CASE DEBRA

This case is interesting, not only in as much in what we know about Debra, but for what we do not know. Consider, for a moment, if the case (see Appendix A) was presented as a set of propositions; all of which were statements, it would appear
something like table 1. The propositions in table 1 were derived by forming the declaration of the programme's question and filling in the answer as an element. (e. g., Does Debra have a fever? No. Therefore Debra does not have a fever.

Debra presents with a stiff neck on flexion, a six hour headache, nausea and vomiting (b); see Table 1. She has a very severe headache (d). She may have meningitis, but she also may have any number of other disorders. Subsequently, as they proceed through the case, the physicians will discover that Debra does not have a fever (e), seizures (f), visual problems (g), nor an abnormal funduscopic (eye) exam (h). She has no cutaneous lesions (cc), nor has an abnormal mental status (k). Her headache was abrupt (h), and she was not recently hospitalized, but was taking an antibiotic (o), and an analgesic. They will consider that her headaches did not have precipitous nor aggravating factors (j), nor has she been knowingly exposed to any contagious diseases (bb). The case progression as illustrated in Table 1, seems to indicate that the physician who dealt with this case returned to physical examination (cc) and history (bb), either indicating rounding out the case, or more likely, an apprehension about the path taken.
**TABLE 1**

**FINDINGS IN PATIENT CASE DEBRA**

(a) The patients name is Debra. Her age is 39 years Her Sex is female, and her Race is BLACK

(b) Her chief complaints are: HEADACHE, STIFF-NECK-ON-FLEXION NAUSEA & VOMITING

(c) Debra has had this kind of headache for 6 HOURS

(d) Debra's headache is very severe (on a scale of 0 to 4 with 0 for very mild and 4 for very severe) she headache is rated 4

(e) Debra does not have a fever.

(f) Debra has not experienced seizures recently.

(g) Debra does not have visual problems.

(h) The onset of Debra's headache is characterize as abrupt (given the choices of abrupt, rather sudden but not abrupt, or gradual)

(i) Debra does not have an abnormal funduscopic exam.

(j) Debra's headaches do not have precipitating or aggravating factors.

(k) Debra does not have an abnormal mental status

(l) Debra does not have a history of polycystic kidney disease.

(m) Debra does not have a family history of polycystic kidney disease.

(n) Debra has taken medications recently.

(o) She has taken ANTIMICROBIALS & ANALGESICS

(p) Debra has not been hospitalized recently

(q) The results of a recent CBC of Debra are known.

(r) Debra's peripheral white count was 13 thousand

(s) Fifty percent of WBC's were PMN's

(t) Twelve percent of the peripheral WBC's which are immature forms:

(u) A lumbar puncture been performed on Debra.

(v) The total WBC count of the CSF is 330/MM³

(w) Fifty-five percent of the WBC's in the CSF were poly's

(x) The protein value (mg/100cc) in the CSF is 635MG/DL

(y) The glucose value (milligram percent) in the CSF is 53MG/DL

(z) The simultaneous blood glucose is 110MG/DL

(aa) The CSF is bloody

(bb) It is not known Debra been exposed to any contagious disease recently (e.g.) meningococcal disease, mumps

(cc) Debra does not have any cutaneous lesions or rash on physical examination
THE DIAGNOSIS

The following summary outlines the diagnosis of the case. In a word, Debra has a case of partially-treated bacterial meningitis which is purulent — (see Table 1). This means that she had been either previously diagnosed and treated (possibly with an antibiotic) for meningitis and had it insufficiently treated, or had some other infection such as a middle ear infection or an upper respiratory infection, which due to insufficient treatment spread to the meninges. The latter seems more tenable, because she has not been hospitalized recently, but there is no evidence of an ear, or upper respiratory infection which may have spread, simply because NEOMYCIN did not ask for it. The antibiotic, which can only be obtained via prescription, for meningitis would conceivably have been given in a higher dose and by intravenous for a meningitis, which would imply administration in a hospital, whereas she would have been, more likely, sent home with antibiotic tablets for something like a middle ear infection. The evidence for a meningitis is her sign of meningeal irritation, which is a stiff-neck-on-flexion; that indicates increased intracranial pressure. From the wording, this would have to have been a sign (a verifiable and objectively obtained fact) attained via physical examination rather than a presenting symptom, because people do not and cannot flex their own necks. As a symptom, a patient may present with neck pain. Other signs of meningeal irritation which might have been looked for, were Brudzinski’s and Kernig’s Sign. The partial treatment of a bacterial illness is suspected because she was previously given antibiotics, yet we do not know precisely when she was given the antibiotics, or why. The laboratory reports, in light of this, seem muddled. If one were to look at her laboratory reports, in lieu of her symptomatology her condition may have been subacute or a chronic brain (pus) empyema, but the protein and the white cells are too elevated; moreover, if we look back at her findings we do not know if she had any focal signs, she does not have a fever, and too short a time line from the symptom for a brain abscess (empyema) to develop. Moreover the
progression of headache for an empyema is more likely to be gradual. The fact that NEOMYCIN does not ask questions about primary infections, and lateraizing and focal signs for empyema, as a space-occupying lesion makes one wonder whether they were unimportant to the case or if were overlooked. Brain swelling would have resulted in confusion and seizures – which she does not have (k, f) and raised blood pressure (of which we do not know). An empyema is not likely to produce as many white cells in the spinal fluid as indicated (v). Nonetheless, it is not impossible that her present condition may have been caused through an improper treatment of a meningeal infection which spread to the brain parenchyma itself, manifesting only recent symptoms, but this would be a rare chronic case, and meningitis would be part of her presenting illness or history.

It is more likely to expect a slight increase in cerebral-spinal fluid protein (x)(around 180MG/DL) and more decreased sugar (y) for a bacterial meningitis. The laboratory tests indicate a bloody spinal tap (aa), suggesting a subarachnoid hæmorrhage which corresponds with the abruptness of the headache onset and the severity of the headache. Yet the symptomatology corresponds with an infection, much more with an intracranial bleeding, leaving one to suspect an unsuccessful tap. Moreover, the fact that bleeds lead to increased intracranial pressure which contraindicates taking a spinal tap, leads one away from such a diagnosis. It may not make a intracranial bleed impossible, but intracranial bleeds are associated with loss of consciousness – which Debra has not had (k), and she has no signs of increased intracranial pressure (l). Although NEOMYCIN did not ask about papilledema, a patient may die of increased pressure before a papilledema emerges. The number of white blood cells in the CSF are at a high level making it less likely for it to be a subarachnoid hæmorrhage (x) probably caused by ruptured berry aneurysm. This is despite the increased protein. Although she has a bloody tap, is is unlikely that there would not be as many white cells (u) and the sugar (y) would not be as low as it was. The
history of the case is very short. Although the characteristics of the headache were attained, there was no search for flu-like symptoms or the possibility of a viral infection. Some absent data were the vital signs, such as blood pressure, respiration, and heart rate. For instance, if one could determine if Debra's blood pressure is normal, it may help rule out intracerebral hemorrhage. Physical exam seemed incomplete. The fact that the headache is only six hours (c) makes the condition acute; the headache was characterized as abrupt (h). The funduscopic exam is normal (j). NEOMYCIN asks some question which are more academic than pathological (m, n). For instance, its queries about polycystic kidney disease are not very relevant to the immediacy of the case, unless the programme is looking for a cause of a systemic hypertension, and ruptured berry aneurysm which may produce a subarachnoid hemorrhage. It did not ask for the blood pressure directly, although it can. In terms of a typical case like this one; one does not think about the cells, protein and glucose in the CSF for the patient is usually being treated in the waiting-room. Tests moreover, are not performed by the physician himself but at the laboratory by bacteriologists and biochemists - in a city hospital. Also the directness of the answers to the question does not allow physicians to examines the subtle fuzzy shadings which some nonquantitative findings may have on a continuum. NEOMYCIN's differential diagnosis definitely points to an infectious process, specifically an acute bacterial meningitis, but it also gives much weight to a subarachnoid hemorrhage. It gives very little weight to chronic meningitis and a ruptured brain aneurysm.

**MENINGITIS**

The class of diseases which are called meningitis are inflammations of the arachnoid mater. This, in turn, affects the CSF in the underlying subarachnoid space by interfering with its flow. Inflammation of the meninges - no matter what its agent, bacterial, viral, parasitic, or fungal commonly produces certain symptoms; such as headache.
The mechanisms that underlie the pathophysiology of meningitis are as such: an infection irritates the meninges producing an inflammatory response. Because the meninges are inflamed they compromise the flow of the cerebrospinal fluid; and this in turn increases the intracranial pressure. Raised intracranial pressure may result in headaches, nausea, vomiting, loss of consciousness, confusion, a bulging or tense fontanel in neonates, and death (Merritt, 1984).

As a result of the inflamed meninges the patient may experience nuchal rigidity (stiff neck on flexion). Observe that the symptoms of neurological disorders are non-specific and thus the signs and symptoms mostly tell the physician what part of the central nervous system is affected by the disorder, but they do not ostensibly reveal the specific pathogenesis of the symptoms and signs. In summary, meningitis and other central nervous system infections are a threat to (at worst) the viability and (at best) the functional integrity of the patient. They demand high level of efficiency on the part of the physician, because patient outcome may be related to delay in diagnosis and treatment by hours and possibly even by minutes. The manifestations are ambiguous, the physician needs to have a high index of suspicion to this diagnosis at all times.
CHAPTER V

METHODOLOGY OF THE STUDY

This chapter has two sections. The first section deals with the methodology of an experiment (experiment I), where subjects perform an semi-interactive parallel diagnosis with NEOMYCIN. This is followed by another set of experiments (experiments IIa, b & c) for an interactive question and answer exchange between the subjects and the programme, and programme exploration. The first two procedures of the second experiment investigate how subjects solve a medical case by asking the programme diagnostic questions. The latter surveys discovery learning and knowledge base design. A second section suggests a method of analysis, using the categories of indices of clausal ellipses with respect to reactions to the information on NEOMYCIN, and a closer inspection of the possible interpretations of a question. The analysis is not an encoding schema for text, but rather a way to select and discuss the nature of the clauses which the subjects used in relation to the task and to their transcript. The goal is to discuss the psychological reasons for the comments physicians have made. An ellipse is repeating the structure of a sentence and its content, but omitting some of the surface expression. This technique is used to tap into the immediate conscious experience, unconscious motivation, or circumstantial influence, which may play parts in the development, maintenance and modification of human behaviour as revealed in one's utterances. A graphic-pictorial analysis is also described.

METHOD

This study has two parts. Both parts are involved making a diagnosis in a medical case. One of the cases which came with NEOMYCIN—Case#1000 (Debra) — was used in the experimental trials. The first part of this study (experiment I) consisted of the on-line presentation of sentence segments, in the sequence of presenting a clinical question, and then presenting the answer to that question related to Debra together
with the next question, and so on. The subjects were free to make comments at any time in the experiment and to ask the programme, at any point, to explain the current question.

The second part of the study (experiment II) has three modus operandis. In all of them, the subjects interact with NEOMYCIN in a manner which consists of typing "keyword" questions (questions about findings) into the programme to which NEOMYCIN replies.

1: In one mode, one of the physicians with greater clinical experience, was introduced to the Debra case without having seen it before. The trial consisted of probing NEOMYCIN for findings and observations after viewing the presenting symptoms, until a diagnosis was made.

2: Five of the physicians, having developed differential diagnoses with respect to the Debra case in experiment I, continued "working" on the case, this time by by probing NEOMYCIN for findings and observations. This experimental condition included all of the physicians, the same subjects who assisted in experiment I, except, of course, the one physician who participated in the previous procedure.

3: Two physicians having done both experiment I and II explore the programme's other facilities to find out more about the medical case.

**EXPERIMENT I**

This experiment takes up the major part of the study. In these sessions subjects carry out parallel diagnostic reasoning as they observe NEOMYCIN solving a medical case. They observe NEOMYCIN asking a question, and then responding with information from file. The pace is regulated by the subjects. In addition to the text presentation of data, the subjects also regulate whether they will continue observing the case as is, or select GUIDON's explanation module to determine why it is asking a particular question. A subject may select 'explain' any number of times for each question, until the programme runs out of explanations for that particular question.
SUBJECTS AND APPARATUS

Seven sub-expert physicians [4 experienced physicians (3 gerontologists and 1 internist) and 3 medical residents] were presented with an on-line medical case presented by NEOMYCIN on a XEROX-1186 Work-Station. The material was presented as a series of questions (diagnostic inquiry) and answers (findings) as differential medical diagnostic procedures. The last items of presented text which the subjects saw on the computer screen were weighted hypotheses in a final diagnosis. (See Appendix A for a transcript of the case Debra). Physicians were asked to think aloud at all times as a method to reveal their general diagnostic reasoning. The resulting verbal protocols were tape-recorded.

PREPARATION

Due to the complexity of this experiment all of the subjects went through a pre-experimental hands-on trial with instructions in order to familiarize them with the mouse, the machine, the programme, and the experimental task. These are the instructions which they were given:

"This is a study about medical diagnoses using a computer programme. You will be asked to operate in a computer environment that is designed to perform medical diagnoses. You will be observing a case that has already been solved by NEOMYCIN. Thus, the programme you will see will both produce and display the diagnostic inquiry and the consequent answers to those questions. It is asked that you read through while 'actively' observing a patient case that is in the process of being diagnosed by the programme and make comments as to your thoughts about both the patient information and the diagnosis, as the case information is being revealed to you. There is no means by which you can use the computer to ask questions, if you are not sure of what you are supposed to do".
PROCEDURE

The short instructions which follow in Table 1, were given to the subjects before they watched a demonstration of the experimental task. The more detailed instructions which followed were wholly integrated into the experimental demonstration (see Appendix B).

<table>
<thead>
<tr>
<th>Table 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SHORT INSTRUCTIONS</strong></td>
</tr>
<tr>
<td>1) &quot;Please read aloud the patient case information; say aloud your thoughts as each question is presented. Indicate what the answer to the previous question tells you about the patient case. Each time a new piece of information comes on the screen, pause to indicate aloud your thinking at that point. What does the present line of inquiry contribute to the case?&quot;</td>
</tr>
<tr>
<td>2) &quot;Please attempt to predict what the next viable question or group of questions might be&quot;.</td>
</tr>
<tr>
<td>3) &quot;Please attempt to predict about what you feel would be a likely response to those questions; predict what some possible answers to these questions might be&quot;.</td>
</tr>
<tr>
<td>4) &quot;Feel free to select, as often as you like and at any time, on the programme menu 'EXPLANATION', before you select 'CONTINUE'. This should explain the rationale for the current question in the CONSULTATION. Please read aloud what it says and think aloud as you make comment about the EXPLANATIONS&quot;.</td>
</tr>
<tr>
<td>5) &quot;If you have chosen 'explanation', Tell the experimenter if the explanation given to you is agreeable, or helpful? If the explanations are not, what alternative reasoning or action should be applied and why?&quot;</td>
</tr>
<tr>
<td>6) &quot;Call up the next question at your own pace, by selecting continue&quot;.</td>
</tr>
<tr>
<td>7) &quot;Indicate the rationale or reason that you think is behind NEOMYCIN's questioning, before selecting EXPLANATION&quot;.</td>
</tr>
<tr>
<td>8) &quot;After an EXPLANATION or EXPLANATIONS you may choose &quot;CONTINUE&quot; to generate the next question. This entire procedure should be followed through to the end of the consultation&quot;.</td>
</tr>
</tbody>
</table>

During the experimental trial the experimenter did not answer questions about the case, only about the confusions of subjects about operations, and he prompted them only if they forgot to think aloud.

The goal of the instructions and demonstration was to prompt the physicians to elicit the following kinds of information:

1) To talk about what they feel were the consequences of the information of the programme's question or answer with respect to the case, and to their differentials

2) Predictions about what viable questions the computer may subsequently generate.
3) Predictions of probable findings that the computer may generate.

4) Making critical reference to the procedure they undergo, including making critiques about:
   a) NEOMYCIN's medical reasoning;
   b) the presentation of the findings;
   c) the hypotheses or differentials;
   and d) the microworld environment.

This probe for information was intended to make subjects to consider their own medical reasoning as well as focus on the programme's "reasoning". The rationale behind this is that the possible differences between the subject's reasoning and that of the programme would be the source of the physician's personification of the programme; e.g. "What does this doctor think he is doing?" There are a number of reasons why a subject and NEOMYCIN would not see eye to eye. NEOMYCIN itself provides a network to indicate these reasons (see figure 1).

The prediction task was aimed at getting the doctors to reveal their present considerations, and the critiquing task was used to reveal subjects' attitudes about elements of the interaction environment. The subjects were urged to state what they were doing at all times. The physicians were never explicitly told to make a diagnosis, just to go through the case and to think aloud. They were not told to make a diagnosis so that it would not seem that their competence was being tested.

THE INTERACTION

When the physicians asked for an explanation of the programme's question, using the interactive explanation environment which NEOMYCIN uses to make an account of its "reasoning" for its diagnostic inquiry, they were asked to read the explanation aloud and

1) To indicate the significance of the information in the explanation environment.
2) To indicate any differences in their reasoning compared to the programme explanations, and

3) To criticise the explanation environment.

The reasoning for these probes are the same as for the diagnostic text.

Transcriptions were made of the audio cassette recordings of the experimental sessions.

The next section describes the methodology used in Experiment II.
Figure 1. Reasons for Differences Between Expert & Expert System NEOMYCIN (1986)
EXPERIMENT II

Within the second part of the study, one subject was free to carry out a diagnostic interview. This trial was designed to see how a physician would behave as a diagnosis was run in this environment from the beginning of the case until the end (without any prior knowledge of the patient) i.e., without having participated in Experiment I (which is similar to what a student would have to do when using GUIDON as a tutor).

Another experimental session, constituting the main part of Experiment II, consisted of the subjects running through an interactive session with NEOMYCIN "entering in" further questions from the keyboard and getting answers, just following completion of Experiment I. All except one subject, who participated in Experiment I, participated in this experiment. The trial acted as an opportunity for the physicians to gather more facts about the patient.

Still another exploratory study was set up where two subjects after performing Experiment I, and the second experiment, were allowed to make a free exploration of NEOMYCIN and GUIDON to investigate the nature of the same medical case and the microworld.

EXPERIMENT II A

SUBJECTS AND EQUIPMENT

The same subjects, using the same equipment as in the first experiment, participated in this part of Experiment II.

PROCEDURE

The physicians were asked if they wished to make further diagnostic inquiries of the programme data base related to the present case. In the second experiment, the physicians could interact with the programme; they could ask questions to obtain findings from NEOMYCIN. They typed in or selected a menu choice of the findings they wished to investigate, and the programme replied.
The physicians could select a finding to "ask" if such a finding and its values were evident in the present case. The physicians were told to think aloud. During the instructions the subjects were, questioned to:

1) Indicate their reason for asking for a particular finding.
2) Indicate the importance of that finding once they had the data, and
3) Critique the environment.

This second experiment was run to see what questions the physicians would ask the programme once they had seen the computer's considerations for its differential in the case. The time allotted to do this experiment, was unlimited, as in Experiment 1.

The subjects worked until they were satisfied. The initial information was presented in the consultation window: it appeared as in Table 3.

Table 3: Presenting Complaint.

<table>
<thead>
<tr>
<th>(Consultation of 21-Oct-85 10: 51: 28)</th>
</tr>
</thead>
<tbody>
<tr>
<td>This patient is a 33 year old woman who has a 6 hour history of stiff neck associated with nausea and vomiting.</td>
</tr>
<tr>
<td>--- Patient – 1000 ---</td>
</tr>
<tr>
<td>The name PATIENT-1000 is Debra</td>
</tr>
<tr>
<td>The age of Debra is 39 years</td>
</tr>
<tr>
<td>The sex of Debra is female</td>
</tr>
<tr>
<td>The race of Debra is black</td>
</tr>
<tr>
<td>The chief complaint of the patient is as follows:</td>
</tr>
<tr>
<td>Debra has a headache</td>
</tr>
<tr>
<td>Debra has a stiff neck on flexion</td>
</tr>
<tr>
<td>Debra has experienced nausea</td>
</tr>
<tr>
<td>Debra has been vomiting</td>
</tr>
<tr>
<td>Next Question? (they can type their choice here)</td>
</tr>
</tbody>
</table>
EXPERIMENT IIB

SUBJECTS AND PROCEDURE

There was one subject in this part of Experiment II. This subject "worked" in an unlimited time frame on the medical case from beginning to end by asking the computer about findings. Again, during the instructions, the physician was told to think aloud, and was questioned to:

1) Indicate the reason for asking for a particular finding,
2) Indicate the importance of the attained finding, and
3) Critique the environment.

Transcripts of the audio sessions were made.

EXPERIMENT IIIC

SUBJECTS AND PROCEDURE

Two physicians, who had participated in Experiments I and IIA had the possibility of exploring the intelligent tutoring system, GUIDON, as much as they wished. In this programme they could inspect an interactive tree diagram and menus representing the programme's diagnostic reasoning for linking findings with hypothesis. The physicians could also check through specific rules in the system and test hypotheses in a number of interactive windows. As in Experiment I, transcripts of the audio sessions were made. Emphasis was placed upon their search in the environment, to discover things about the patient Debra, her condition, and other facts, and NEOMYCIN's treatment of rules about relevant disease conditions. In this computer environment mode, physicians can look at its many aspects.

METHODS OF ANALYSIS

The following section describes indices used to specify types of clausal ellipses (e.g., Karen ran for president and Bill for mayor) in text cohesion. Where there is an ellipse there is a presupposition of a preceding item — which serves as the source of the missing information— in this case running. It is anticipated that ellipses, which are
normally an anaphoric relation—referring to preceding text, will be a source of presuppositions for the subjects when they have to make judgments under uncertainty.

However, as pointed out by Stubbs (1983), such things are features of surface syntax; e.g., Mary was president, as was Jack before her. The "object" of the inference could be the surface of the sentence itself; What was Jack? Jack was president. When asked a WH-question, there is an unknown quantity X- (where, when, who, whose, which, what, how), and Yes/No questions form part of the answer in their form. The type of WH-question plays a part in establishing the possible answers to the question. However, answering a question is more than a lexical matching task. Thus, in this study, only the labels of textual cohesion are used as a way to demonstrate how a computer displayed text and how the subjects' responses tie into a cohesive whole. Keeping a situation coherent is a form of problem-solving in itself; not only in understanding the meaning, but having it make sense. As the text in this study is the only place the subjects will get their information how much of their assumptions will come from the textual cues, and how much from the use of their imagination (guided by their intellect) to fill in information to which they do not have access. Since they learned all they know about medicine and linguistic interchange from their contact with reality, it can safely be assumed that their assumptions will reflect their mental representations of the day to day encounter with patients and other member of the medical community. The information sought, in this study, therefore, is not how subjects omit, or substitute words in their sentences, but how they interpret non-ambiguous, and non-vague text, not in a purely textual context, but in a situational one. When encountering new information, subjects may assume that in the structure of the series of textual strings that something is supplied or understood when they feel that they have a gap in their information. Therefore this method of analysis concentrates on types of responses to text which imply that the subject has made a presupposition. It is followed by a concentration on the possible interpretations of NEOMYCIN's
diagnostic questions, and the definition and description of the 5 major salient categories which emerged from the verbal data: reasoning, explanation, predicting, critiquing, and anthropomorphisation.

The think aloud protocols should reveal the subjects' general diagnostic reasoning (see figure 2). Figures 2 and 3 show how the probes tap reasoning, explanation, prediction, and 'critiquing'. These probes which the subjects are given, incite them to predict questions and findings. It should also incite them to make a critique and to explain the significance of presented information, with respect to information provided by the findings, to the questions and to GUIDON's explanations. The same holds true for the suspected effect of the experimental probes for system findings as well as questions (system inquiry), save that, for questions (see figure 3), the subjects may encounter voluntarily related computer explanations. Perhaps, it would better represent the purpose of the "probes" by drawing the arrow that leads into the probe box as coming from the "general diagnostic process/procedure" — which refers to their basic medical reasoning. This would be true for both figure 2 and for figure 3, as the probes are to tap into the physicians' reasoning process. In these tree diagrams, the line leads from the stimuli boxes (i.e., the system findings and inquiry) to denote that the probes are asking for something more than the generalized medical reasoning. Moreover, the subjects are, in a sense, probed to respond to system explanations (figure 3). They are expected to choose explanations freely and, as a matter of course, will be in a situation of making inquiries of the system and, therefore, will not really need to be specifically probed for commentary; they simply need to continue thinking aloud.

The critiquing process should communicate a meta-level of the diagnosis. That is, commentary about the diagnosis itself. While the general diagnostic procedure addresses the reasoning, critiquing should tap into how the case is being solved per se, as well as dealing with subjective reactions to the experimental media. The indices used
In this study to classify text in subject protocols refer to any responses to verbal or written material.
Figure 2. Facets of Experimental Probes with respect to subjects' responses to findings presented by NEOMYCIN
Figure 3. Facets of Experimental Probes: Subjects' responses to QUESTIONS & EXPLANATION presented by NEOMYCIN
THE EFFECT OF QUESTIONS: LINGUISTIC PRAGMATICS

The interpretation of received information and human action can be discussed in semiotic terms. In human/machine interaction, the transfer of information from the human being to the machine, takes place via some action on the system, which is intended to change some state (Rasmussen, 1986). Subject attention, on a system like NEOMYCIN, requires some attention be aimed at surface level programme functioning and within the content of the microworld domain. Shank & Abelson (1977) propose that people have scripts of situations, and they discuss how people selectively focus in texts on the main elements which they need to reach goals, while ignoring minor elements.

Thus, a study of semantics would focus on what the subjects understood about the case, in light of what the linguistic terms meant. The analysis is intended, on the other hand, to capture the interpretations the subjects made, not exclusively with respect to connotative or denotative meaning in reference to the literal understanding of the presented text, but rather to assumptions made about the extra-linguistic situation. Used to denote the subjects verbal responses to data, the analysis, in part, consists of a modification of Halliday & Hasan's (1976) indices of types of rejoinders:

The "question and answer" sequence is a standard pattern in language. At the same time there are other sequences involving rejoinders of one kind and another. Not all questions have an answer; but no less significant is the fact that not all answers have a question. Any observation by one speaker, whether it is a question or not, may be followed by an observation by another speaker that is related to it by some cohesive tie. We shall refer to this very general category of sequel as a REJOINER. A rejoinder is any utterance which immediately follows an utterance by a different speaker and is cohesively related to it (Halliday & Hasan, 1976, p. 206).

Their taxonomy was chosen because it can address the subjects' responses to questions as well as statements (findings, explanations) and commands, like probes.
Among the possible interfaces are the computer-computer interface, the subject-computer interface, and the subject-experimenter interfaces. There was, by design, almost no subject-experimenter interface after the initial instructions. The computer-computer interface describes the programme's responses to the programme's questions; e.g., Q. How long has Debra had this kind of headache? A. Six hours. There were two types of subject-computer interface 1) The subjects' questions about the computer's questions, explanations, or findings, in which they actively used the programme's explanation facility to make inquiries—to ask the programme questions and 2) the subjects' verbal commentary about the computer's questions, explanations, and findings.

Typically, in an experiment such as this, as the programme asks a question, the subjects expected that the data base should produce a veracious, appropriate or direct response to a yes/no question or to a WH-question, even if it is to say, "I do not know" (an appropriate indirect response). A direct answer is a sufficient answer. It fulfils all of the Gricean Maxims, direct answers ought to be informative, truthful, relevant, and parsimonious (Grice, 1975). It is one which answers the intended question of the questioner—which would be accepted as solving his or her problem. Moreover the answer doesn't answer more that what the questioner intends (Åqvist, 1975), which is in accord with the Gricean Maxim of Quantity. Appropriate, sincere questions and direct responses should not yield verbal retorts from the subjects, which would indicate that they are unsure about the nature of the interface. Perceived inappropriateness will make the subject question the rationality of the programme interface. This holds true for the subject-computer interface, as well.

**WH - & YES/NO QUESTIONS**

At first sight, the answer to a WH-question seems very different from a yes/no question. The simplest answer to a WH-QUESTION, however, is to fill in the blank; e.g., "What time is it? "7:30." The nature of the response to a question is, however, a
function of how it is interpreted by the subject. Whether a subject perceives the object of the question to be asking for information or requesting that some behaviour be performed is a way of differentiating the questions presented by NEOMYCIN. For instance, in everyday conversation if one were to ask the similar question, “Do you know what time it is?”, “Yes.” would be a rather unexpected reply since, by such a question, one is usually being requested to tell the time, if one knows it, or being chastised for being tardy, in which case it is a rhetorical question.

The answer to the information which a sincere yes/no question calls for specifies the polarity and presupposes all else, by making explicit just one thing and leaving the rest to be presupposed by the ellipse; e.g., “Do you think that this corrupt administration, which has lied to the public and continues to undermine justice should remain in office?” “Yes.” Similarly, many things can be implied in a WH-question, e.g., “How many dreadful years have you been married to this insipid little jerk?” “Five.”

The WH-Expression itself indicates whether the missing item is participant (Who left the TV on?) or circumstance (Where is the tie?). Also, if the expression is participant, whether it is from a limited set (which one), whether it is human (who), or non-human (what), whether the question is one of degree (how much) or kind (what sort) is indicated by the WH-expression itself. The same holds true if the expression is a circumstance; e.g., if it is time (when), or place (where), cause (why), or manner (how), etc. (Halliday & Hasan, 1976). The WH-concept is an unknown quantity which indicate action, things, or events (Stubbs, 1973).

Despite this similitude between yes/no questions and WH-questions, yes/no questions will generally have vacuous presuppositions. They may be the disjunctive (either/or) of their possible answers (e.g., Is it raining?: Either ‘It is’ or ‘It is not’ (Levinson, 1983).

Alternating questions have non-vacuous disjunctives (Is it this or that?). For instance if one asks, “Did Fred or Billy speak to Chuck?” an affirmative may
presuppose an inclusive or, where as either/or presupposes an exclusive or (Åqvist, 1975).

WH-questions introduce the presuppositions obtained by replacing the WH-word by the appropriate existential quantified variable. Theses presuppositions are not invariant to negation. Stating something, especially a command in the form of a negation, semantically leaves wide open possibilities for what may not be prohibited (Levinson, 1983). Figure 4. shows Halliday & Hasan's (1976) types of rejoinders.
Figure 4. Types of Rejoinders from Halliday & Hasan (1976, p207)
QUESTIONS AND OTHER THINGS

The subjects are often making comments about the computer questions. Commentary about the computer questions can be treated as indirect responses. This is so, because responding to a question depends upon differentiating whether the intended object of the string is asking for information or requesting that an act be performed. For example, if one says "Don't you think it is cold in here?" if the intent asking for information the answer "Yes" would be fine. However, if it is an indirect request, it would be semantically congruent to the connotatively loaded string, "Why don't you close the window?" = "Close the window; it is cold in here.

Moreover, the subjects are making commentary about things which are not questions, but findings and computer explanations. Indirect response is the one kind of response which is not an answer in the defined sense.

Examples of indirect responses are commentary, disclaimer, and supplementary response (Halliday & Hasan, 1976).

Commentary refers to the speakers attitude to the question, which may connote acceptance, rejection, or confusion. They are respectively denoted here in this example: "Has the patient ever been inoculated for malaria?" "Of course!" — "Don't be ridiculous. "— "I don't know why you are asking this. "

A disclaimer, displays rejection, involves moving from a yes/no question to a WH-CONTEXT, or vice-versa. e.g., "Is the vile ready" "Whose?"; "What?"

Supplementary responses, are responses which gives information other than is asked. Supplementary responses are usually associated with yes/no questions is answering by implication e.g., "Can you read the sign?" "If you get out of my light. " (Halliday & Hasan, 1976).

These categories are useful because the yes/no questions and the unknown quantity of the WH-QUESTIONS were often weighted by degrees of certitude which can be handled by modal logic; e.g., She has a good chance of having a fever.
Halliday (1985) shows examples of expressions of probability:
1) I think...
2) I'm certain...
3) It is likely...

as well as expressions showing a sense of implication
1) It looks like...
2) This fits in with...

Halliday (1985) refers to the psychological subject of topic as the theme and the comments around it the rheme, as well as the actor who are the logical subject.

His goal is to show the relations between clauses which can be accomplished through different sorts of expressions. Some expressions are for elaborating; for instance-exemplify: "It is day time". "It is noon". Some expressions are means to clarify; i.e., "She appeared perplexed". "She was thinking about the rendezvous".

The second clause elaborates upon the first clause by restating it, changing it, refining it, or adding new elements and attributes; such as - The clock doesn't go; it's broken. Expressions which typically add something new to a previous utterance can use links like - and, but, nor, instead, or except.

Some expression enhance or qualify a clause by making reference to time, space, manner (i.e., comparison or means) and casual conditions (i.e., reason, purpose, positive condition, negative condition, or concessive condition. A concessive condition is: if \( p \) then contrary is expected of \( q \) (Halliday, 1985).

As mentioned above, the subjects did not only react to questions, but to the findings and explanation, and to the probes, which were taken as commands, as well.

Halliday and Hasan, (1976) indicate that if someone responds to a statement their rejoinder may either indicate their assent (acceptance), for example, "The contract is in order". "Of course it is", or their rejection where they respond by contradiction: "It is not". A statement or a command may by responded to by a question rejoinder: querying it or eliciting supplementary information. Reactions to
command may display consent "Pick up some milk on the way home" "I shall" or refusal I shall not".

**PORTRAIT OF A QUESTION**

Since phrasing can colour the interpretation of information, let us take a look at how one question produced by the programme, and how its quasi-intentional face affected the subjects.

The yes/no Question (14) seems to have the force of a WH-QUESTION. Question (14) *Has Debra taken medications recently?* may be interpreted as follows:

(14i) Has Debra taken medications?
(14ii) Does Debra take medications?
(14iii) Has Debra been on medications recently?
(14iv) Is Debra on medication? Now?

All of the instances could be answered as though it were a yes/no question, but the nature of a medical inquiry brings the question into context. If the sentence is taken for its role in the diagnosis, it may beg one of several corollaries:

[Has Debra taken medications recently? And if so what medications?]
[Has Debra taken medications recently? For what ailment (Or why? For what purpose?)
[Has Debra taken medications recently? What type of medications?]

Therefore, a question formulated like

(14v) If Debra has taken medications recently, what medications is she taking?

is a reasonable expectation if one is expecting the machine to answer the question. It could also be the other way around it could be an indirect request for information, and be in essence a command:

(14vi) Please indicate what medications Debra has taken recently?

If the subject’s answer is in the form of a consent he or she might feel that the question is either:
(14vii) With respect to Debra's condition at this point what are the medications that she might be taking to alleviate her symptoms? (Self-administered or prescribed)

or

(14viii) With respect to Debra's condition at this point what are the medications that she may have been given by a physician to alleviate the symptoms or cure the disease?

The subjects might feel that one is supposed to interpret ((14vii- Has Debra been on medications recently) + Guess what possible kinds of medications these might be? That is to interpreted the interaction as an experimental test of their medical competence and take it as an implied command. Another similar interpretation of the string as an indirect command-like response would interpret the string to mean:

(14ix) In light of the patient's present condition what substances could cause all or some of these symptoms.

The answer provided by NEOMYCIN -{a}: (14a) "Yes", in this case, would have been inappropriate in that it would not be complete. The physicians expect to find the answer in the explanation facility. Therefore the sentence string, Has Debra taken any medication (where D=patient Debra, & X=medications) -((I ASK you whether D) has taken any X) could be answered as two strings e.g., ((I request you(you Tell me whether D) has taken any X) - (and if so (I request you(you Tell me what D) ?X)). Even this has some fuzzy elements; who is the "I" for instance? Who is the "you", which could be the consulting physician, the database, the patient or the programme. X being only slightly more clear is always considered prescription, over the counter, but also includes street drugs. So although it is a yes/no question all of the subject treated is as a WH-question, and proceeded to guess what type of drug the patient might have taken (Gazdar, 1979).

The subjects protocols consisted of reactions to the experimental probes, the programme's questions, its answers (findings) and its explanations.
COGNITIVE OPERATIONS

Five important categories emerged from the verbal data which characterized them. Namely, they were reasoning, explanation, predicting, critiquing, and anthropomorphization. *Reasoning* is used here to refer to the activity of thinking and analyzing to draw inferences and conclusions in order to make phenomena: assertions, ideas and beliefs intelligible. In this study the reasoning focused mainly on the medical clinical reasoning with respect to solving the case. The term *explanation* is used to denote the explicit verbalizations which are intended give reasons to justify and account for actions and events by bringing forth implicit considerations which the subject used to interpret the information. It is the conscious declarative explications of their actions and considerations. *Prediction* refers to talking about what one feels the programme will do next or how the patient's condition or the course of the patient's alleged illness will unfold. *Critiqu* is to pass judgement, review and evaluate the merits and faults, qualities and characteristics of a phenomena. *Anthropomorphisation* is to personify; attribute human qualities and personality to things, especially faculties of intelligence and purposeful free will with respect to intentional actions and rational concerns.

PICTORIAL ANALYSES

In the results there are three types of graphic-conceptual representations of the subjects' medical reasoning. The first kind is termed a schemata. It follows a convention which is described in detail in the results chapter. The function of the schema is to indicate the differential diagnosis and the corollary medical consideration which were made by the subject at each point of a given textual stimulus (be it a question posed by NEOMYCIN, a finding, or an explanation). Chronological order runs ahead from the top of the graphic to the bottom. Arrows are used to facilitate reference between the observation items: findings, computer questions and explanations, and the differentials and considerations. This analysis was performed
on one subject, focusing on the observations which were based on history and physical examination. The reason for this is that physicians are taught to make diagnoses and to even (if necessary) start treatment before they ask for confirmatory laboratory tests. Of this same ilk, is a representation of the doctor who asked diagnostic questions to NEOMYCIN from the onset of the medical case. The categories differ here. After the observation, the subject states an **intent** to perform a diagnostic act, then performs that **act** (types a question into the programme), receives a **finding** and then indicated the **significance** of that finding.

The second sort of diagrams are called network schemata, and they represent mainly the subjects' reaction to findings and explanations in temporal order. As the diagrams above, the network shows the observation, the doctor's differentials, as well as their medical considerations. However the differentials in these diagrams are the focus, and thus they are represented with strong lines to indicate what is happening with the differential as the subjects proceeded through question #7 in this study.

The third variety are reasoning considerations/deliberation tree diagrams which for the one doctor it was applied to simply indicates the medical and extram edical commentary made at the discovery of each new observation.
CHAPTER VI
RESULTS AND DISCUSSION

This chapter will present the findings of the study. It has two major sections, one pertaining to the results of the first experiment, and one describing the results of the second set of experiments.

Within the first division are the results of experiment I, focusing initially on an experienced internist: subject 6. There is a series of examinations of 5 major categories of subjects' actions, in reference to the information they encountered: reasoning, explanation, predicting, critiquing, and anthropomorphisation. Excerpts of the best examples from other subjects are examined as well. The next section looks at the results of the indices of clausal ellipses, with respect to reactions to the information on NEOMYCIN. After these general analyses there are detailed case studies. The main case study is of an experienced subject (a gerontologist) who displayed the most conspicuous examples of critiquing and anthropomorphisation: subject 4. Following that, is a comparison between the experienced physicians (subjects 1, 4 & 6), with respect to the less experienced medical residents (subject 2, 3 & 7), which concludes with a focus on the general reasoning of another gerontologist: subject #1.

In the second division of the results, there is an overview of the outcome of experiment II. There are three sections, respectively. The first describes the behaviour of the subjects as they interacted with NEOMYCIN, following their interaction with it in experiment I. This is experiment IIA. The second section shows the experimental results of subject 5 (gerontologist) who ran a complete medical investigation of the case by asking NEOMYCIN diagnostic questions. — experiment IIB.

Finally, the third section takes a look at some of the subjects free exploration around the microworld and GUIDON — experiment IIC.
Below is an outline of the participants and their involvement in the four experimental trails.

**TABLE 4.**

OVERVIEW OF SUBJECTS

<table>
<thead>
<tr>
<th>SUBJECT</th>
<th>BACKGROUND</th>
<th>EXP I</th>
<th>EXP IIA</th>
<th>EXP IIB</th>
<th>EXP IIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject #1</td>
<td>Gerontologist</td>
<td>√</td>
<td></td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Subject #2</td>
<td>Early Resident</td>
<td>√</td>
<td></td>
<td>√</td>
<td></td>
</tr>
<tr>
<td>Subject #3</td>
<td>Late Resident</td>
<td>√</td>
<td></td>
<td>√</td>
<td></td>
</tr>
<tr>
<td>Subject #4</td>
<td>Gerontologist</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subject #5</td>
<td>Gerontologist</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subject #6</td>
<td>Internist</td>
<td>√</td>
<td></td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Subject #7</td>
<td>Late Resident</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**EXPERIMENT I**

Experiment I placed the subjects in a situation where they exerted a moderate amount of problem-solving effort. Although they were not asked to diagnose the case, they all tried to diagnose Debra’s illness as the information unfolded on the computer screen. The presented text was in the form of either questions, findings, or explanations of questions.

**REASONING**

Reasoning refers to the medical considerations which the subjects made in trying to solve the case. The subjects thought-aloud about what they saw on the computer screen, and went about solving the medical case, relating findings, symptoms and signs to hypotheses, disease conditions, disorders and the differential.
observing the subjects’ responses to these textual stimuli, it was noted if they discussed content which were medicine-specific, as subject 6 did here in excerpt 1:

| excerpt 1 | Subject 6 [SUBJECT REASONING A low-grade fever with a stiff neck and a bad headache, nausea and vomiting...One would be more inclined to think of something vascular in the brain. Something ahm - ilike a hemorrhage. A high fever with those symptoms would more - more make me more inclined to think of infections.] |

Note how the physician keeps the differential broad and vague by using broad categories and by utilizing stochastic language — "more inclined". The experienced physician (subject 6) will be examined in this section of the results, as this is the physician who is the most verbally demonstrative of medical reasoning. This physician was not more or less verbose than the other subjects overall. These results, of this subject, were in part due to this physician feeling that the study was designed to tell him things about his own medical reasoning. It is interesting to note that the physician is not interfacing with the machine as though it were merely a 'spreadsheet' on Debra, but rather in the way a doctor asks questions, about, and to a patient he is examining, or as one doctor speaks to another on-line about a patient they are co-examining. Co-examination is a consultation where one physician is in contact with the patient or the records, and the other is on the telephone. The investigation speaks of a process of diagnosing Debra’s problem quickly and efficiently and is not academically motivated.

One observes below (excerpt 2) how the subject is talking about the running of the differential diagnosis as though it were an interview in an medical office. The direct reading of text off the computer screen is marked in italics.
Debra: Patient is thirty nine years old female black, ah chief complaint headaches, stiff, neck on flexion nausea and vomiting.

Computer: How long has Debra had this kind of headache? 

I'm going to continue, I guess. And Debra has a head ... and she has stiff neck on flexion. She complains of nausea and she complains of vomiting. Ah... and the next question: ...How long has she has this kind of kind of headache? ah. I would like to know the nature of the headache; whether it was a unilateral or a bilateral headache or all over the head. Ahm Ahm, the severity of the headache; whether the headache was associated with the vomiting, ahm Her neck was stiff? How long it's been stiff? And ...Has she had any types of headaches like this before; has she had any fever? How high is the fever? What is the nature of the fever? Is it a ...This it ...Is it a fever that ah... abates and then exacerbates and abates and exacerbates? In other words is it a fever that we call a ...[inaudible] fever or not?

This is a part of a larger passage; the subject continues after this to pursue findings for some time (see figure 10). All the physicians suspected that Debra had an infection: namely meningitis, deduced from the presenting symptoms with a high probability of intracranial bleeding, which they considered from the abruptness of the headache.

All of the aforementioned questions, in excerpt 2, revolved around history taking in the clinical interview — the exception being fever, which is determined upon exam. The nature of the headache, and the stiff neck, and their effects on her central nervous system, especially in terms of infection had concerned this physician, but the doctor also asked about family history of hypertensive headaches, taking into account her race. Upon learning that the headache was very severe, the transcript indicates a continuation in looking for evidence of infection. The subject begins here, in excerpt 3, to make links to states and conditions, such as increased intracranial pressure, which are projectile vomiting, and seizures The inquiry, again, is one of basic history, such as, exposure to infection, but with interspersed examination questions.


By asking about her work, her contact with children and the public the subject is making an elaboration of the same question, which explicates in the last sentence of the excerpt where the physician wants to know if where she may have got an infection. As we have seen, the subject was not severely hampered by the experimental interface, in terms of forming a diagnosis. As all the subjects did, this subject went along diagnosing the case from the available information provided by the findings. All the physicians talked about the nature of the clinical task and their medical reasoning pertained principally to the findings. The other categories are directly linked to reasoning. They are a product of reasoning, but the focus is more refined.

EXPLANATION

There are instances where the subjects explain their actions, per se. Again, one can observe the clinical reasoning of Subject 6 followed by an instance of explanation.

By indicating that race motivated the question, the subject is making a clarification. The focus of the subject’s explanations were about the subjects’ reasoning. The considerations of hereditary causes for the headache is made more
explicit as the subject indicates the reason for the considerations made. When the subject elicited a rationale for their medical reasoning it provided more refined information about the nature of the decisions which were taken at each step of the differential diagnosis. In the section, addressing the results on anthropomorphisation, (p.103), it is evident, as well, that the subjects make rationalizations, not only about their own action, but about those of the programme. This gives one an illustration that the subjects are conscious and attending to two sets of intentions, that of the programme's and that of their own. However, even if personification is occurring, the material considerations of subjects' explanation may be entirely medical, rather than conversational, as shown below from an excerpt of the experienced gerontologist — subject 1:

| excerpt 5 | subject 1 | [SUBJECT REASONING] I still don't know why we're asking about polycystic kidney disease. Ah unless, unless... ah people with polycystic kidney disease do have a higher incidence of hypertension. [SUBJECT EXPLANATION] Hypertension can be associated with cerebral bleeding [disease] so, maybe, that's what they're looking for here. |

Here we have an instance where it is assumed that the programme must be performing a step for a useful reason. The subject elaborates on this condition indicating that what is in his/her own differential must somehow be associated with the current path.

Subject 6 shows how closely subject explanation is tied to reasoning. Indeed all of the previous excerpts which this doctor verbally (as opposed to typing in requests as in experiment II) asked diagnostic question, or rather, stated a desire to do so, begs an explanation of their intentions to fully understand their clinical significance. Below, this subject's actions and intentions were made explicit:
Subject's 6 protocols are indicative of a subject in which most of the medical consideration were oriented towards self-reasoning, in light of the case, rather than in personifying the interface.

The results of all subjects' explanations are mainly tied to findings and questions. Explanations did not stimulate SUBJECT EXPLANATION, as the reasoning already had been explained by the programme. If, on the other hand, the COMPUTER EXPLANATIONS were vague or ambiguous, this then lead to the subjects presupposing what the system meant (see excerpt 7) by inferring from its previous performance. These presuppositions are oriented in light of the type of uncertainty, which the physician is normally accustomed — an instance of anthropomorphism. Below in excerpt 7, the computer explanations are underlined:

The subject's reasoning about the programme's explanation is in accord with the Gricean Maxims of *quality* (truthfulness) and *relations* (relevancy), with respect to the patient data. The subject then temporarily questioned the veracity of the message and thus reduced the certitude of the differential constructed. The SUBJECT EXPLANATIONS, occurred as a response to the experimental probes. The probes which asked the subjects to indicate the importance of new information (and to tell why he or she has that opinion) proved to be good way to get to know the rationale behind the reasoning of the subjects in this experiment.

95
PREDICTION

In terms of making predictions, the subjects made predictions in reference to new findings and questions. However, most of the subjects did not frequently respond to this probe directly (as the medical resident [subject 2] did) in excerpt 8 below.

| excerpt 8 | subject 2 (REASONING) If she'd had these headaches before; is she taken any medications for them? What does she do to make it better? If she had then in the past, is there a family history of this? These are questions that I'm going to ask, that I predict that you are going to ask. |

Even in this instance, one sees that this probe served to stimulate the physician to talk about the current medical picture. The probe which asked them to predict future questions and answers lead them to feel as though they were supposed to guess findings for strings which could look as though they were instructions, such as:

14) Has Debra taken medications recently?
   Yes:
15) Please elaborate.

It would have not be so curious if the subjects simple responded to "15) Please elaborate." as a command, which they did. Although how can one elaborate about what does not know, except to guess? What is more striking, however, is that the subjects attempt to answer question 14) as though it were a command. The subjects tried to predict what medications the patient is on.

The subjects made predictions concerning the diagnostic inquiry which NEOMYCIN has just made. Here is an example made by subject 3, a medical resident, making a prediction with reference to the programme's question 7) Does Debra have visual problems? This subject interpreted the question as being an investigation as to whether or not Debra has photophobia: "I think she is going to be photophobic either because she has a migraine or either because she has meningitis".

By making predictions about future questions, as in this instance of the medical resident subject 7, 'I think it is going to ask questions to determine if she has a bacterial
infection, "the subjects were trying to making assumptions about the programme's line of reasoning. Therefore, monitoring subject reasoning via prediction allows one to get at a representation of the subject, as well as the subject's representation of the programme and the case at a given time. The subjects, however, did not consistently try to make predictions after each and every question, which the exception of subject 7, possibly due to the unnaturalness of the task, which would interfere with their medical reasoning. It was thought unlikely that any particular question in vacuo could clearly presuppose a finding. Yet, the influence of NEOMYCIN's questions provide sometime subtle suggestions to the subjects if they had not considered a particular avenue.

With respect to the hypothesis of active versus passive participation in this study, the subjects were active. There was a possibility that the subjects would have inferred, from the nature of the programme's questions, an agenda which they felt compelled to follow; and thus, adopt a 'wait and see' attitude. The subjects did, however, pursue their own hypotheses and tracks of reasoning, despite the fact of not being able to necessarily attain the information which they verbally queried about. No overt or covert "suggestion" presumed by the subjects as to the nature of the programme's diagnostic reasoning produced, such an overt response on behalf of the subjects as to make the interaction passive, like watching a drama unfold on television. The acts of predicting the answers to the questions revealed what the subjects felt about the effect of knowing the answers of the questions posed by NEOMYCIN. With reference to the programme asking Debra if she has ever had exanthema, subject 7- (a medical resident) made this comment when the programme displayed that the finding was "Unknown" - "Unknown. That's normal. People never know if their kids have mumps or not. " Curiously, this subject did not say that people do not remember if they, themselves, had chicken pox or measles in reference to themselves, but, rather, if their children had it. Since the case has nothing to do with the patient's offspring, but the patient herself,
this prediction is clearly tied to a clinically-strategic generalization about how patients answer questions.

Another aspect of prediction, which was unanticipated by the experimenter, emerged. In instances of uncertainty, the subjects made a predictive assumption about the purposive reasoning of the programme, especially in terms of how it would explain itself. This, in part, determined how they assumed their queries, when asking for an "explanation", would be addressed by the machine. So it seems that their predictions were tied in with personifying the intent of the programme.

CRITIQUING

Critiquing is a special form of an explanation; i.e., a judgement. As it is for subjects' explanations, the critiques of the subjects were being aimed at the programme's questions and their associated computer explanations. It is subject 4, an experienced gerontologist, who demonstrated the most florid examples of critiquing and anthropomorphisation.

When critiquing occurs, it is focused on the programme's question and is principally followed by a request for a computer explanation. In effect, these interactions may be summed up by the statement made by one physician → "I'm going to ask them to explain this because I don't know why they're asking it". Using the computer explanation facility is a procedure as straightforward as choosing "explanation" from the programme's menu instead of "continue". Hence, a more unambiguous, specific question could not be asked. The main observation that can be made is that beyond any medical reasoning, when a subject selects EXPLAIN instead of CONTINUE, he or she is asking the programme some sort of question.

We have previously seen an instance of a programme EXPLANATION underlined in excerpt 7. Here follows another example to see how they are formulated. They are displayed in the form of a statement of a question and the answer.

• *Why are we asking for simultaneous blood glucose?*
We are trying to determine whether the CSF Glucose measurement was normal.

Normal CSF Glucose is a finding whose value needs to be determined.

The subjects' response to the programme's explanations depended upon their expectations, beliefs, desires, thoughts wishes and intentions, and this governs the response to what is consciously known or unconsciously recognized. Thus, it was interesting to note what was the force of the question the subjects intended to put forth, upon selecting explain and to what is the programme's faithful reply was aimed at. If the interface produces a: "That is not the sort of answer I was expecting", or "That's not what I want now" type of interaction, then the interface illustrates a situation where the machine does not do what the subject thinks he or she is telling it to do. The results below show how interpretation affected reasoning.

Here the reply to "explain" is appropriate, with respect to the intended question of subject 4.

<table>
<thead>
<tr>
<th>Excerpt 9</th>
<th>(Computer Question)</th>
<th>Does Debra have Problems?</th>
</tr>
</thead>
<tbody>
<tr>
<td>[Subject Critiquing]</td>
<td>So she hadn't had seizures. So he wants to know if she has visual problems, and he's probably asking that because he's considering migraines. She doesn't have a fever. She doesn't have seizures. She just has a very severe headache which won't go away with a lot of other side effects. So he's thinking. He's doing two things... Is she got visual problems?...He's wondering if it's migraine as a symptom of migraine, or could the visual problems be an exhibit of seizures. It could be seizure that she doesn't call a seizure. That could be possible, or she could have a pituitary tumour pressing on her optic chiasm and that would cause visual problems. That's what I think. Well let's see if he agrees with that...</td>
<td></td>
</tr>
<tr>
<td>[Subject Prediction]</td>
<td>If that's he asks that next</td>
<td></td>
</tr>
</tbody>
</table>

The subject does not necessarily project self-reasoning onto the programme, but does replace the monitoring with her own. The consideration of migraine is due to the situation, and two things, that the physician feels makes the inquiry viable, are tied to being curious about visual problems migraine and seizures. There is nothing to indicate until this point, whether or not the subject would have asked this question at this time. As well, the label of the symptom given by the patient is thought to be
incongruent with a professional physician – a seizure that she does not call a seizure.

The subject then selects "explain".

excerpt 9 continued  
<table>
<thead>
<tr>
<th>COMPUTER EXPLANATION</th>
<th>[i.e. WHY are we asking whether Debra has visual problems?]</th>
</tr>
</thead>
<tbody>
<tr>
<td>COMPUTER EXPLANATION</td>
<td>We are trying to determine whether Debra has photophobia. Photophobia is more specific than the finding we are asking about.</td>
</tr>
<tr>
<td>SUBJECT CRITIQUING</td>
<td>I agree, but there are other visual problems that would be interesting to know about, too. She might have a photophobia, and still have other visual problems. So maybe we will see if they explains a little bit more about that.</td>
</tr>
</tbody>
</table>

This is almost in the form of a debate or conversation; it becomes clearer that the physician is not in concurrence with the reasoning behind the programme’s question, but not whether the question was ever thought to be unsuitable. This was an example of a prediction being tied to 'critiquing'. The subject interpreted what the "doctor" is thinking, and assumed what the resultant behaviour would be. In contrast, the intent of subject 2 below in excerpt 6, was to clarify a finding, not the reasoning.

Recall that NEOMYCIN does not answer questions about findings.

excerpt 10  
| COMPUTER QUESTION | Has Debra has been taking medications recently? |
| SUBJECT CRITIQUE  | Very basic important question. |
| COMPUTER ANSWER   | **Yes. |
| SUBJECT EXPLANATION | I'm trying to think of which medication can cause this. There are certain street drugs or certain type of medications that raise your blood pressure and that could lead to intracranial bleeding; I worried about that too, there is a lot of speeds, pills, and even. Explain that, let's see. |
| COMPUTER EXPLANATION | We are trying to give thorough consideration to any recent finding. Recent medication is a new finding that needs to be clarified. |
| SUBJECT QUESTION  | Does she give the medication—I mean when we ask her for the medications? |

This could be considered a task consideration where the subject queries about what the explanation facility provides. The weight of the medication question was requesting: i.e., "Explain to me what medications Debra is taking. " NEOMYCIN is answering the question, "Why am I asking Debra whether she is on medications? " This consideration which the programme takes is more oriented towards it goal of teaching
clinical reasoning, by way of getting the student to think about strategies of clinical reasoning.

This subject assumed (as did most of the other subjects) who asked for an explanation at this point assumed, that the programme was going to reveal the medications. Again, this could be an artifact of the subjects perceiving a test environment. Only one subject was surprised to find that the answer to the next question on NEOMYCIN revealed the general type of medications the patient was taking. The purposive intention of the programme was at question, since it asked questions which could have be deemed academic instead of questions which needed to have been asked quickly in an emergency situation. Recall, that if Debra has meningitis she could be dead within a couple of hours without treatment. Subject number 4 criticizes the “doctor” by indicating that there is an easy and quicker to get at what is being implied by the programme’s explanations, since NEOMYCIN replied to a question too generally for the clinician’s taste.

| excerpt 11 | SUBJECT 4 | COMPUTER QUESTION | Many toxicities could give a headache. Also street drugs would be interesting to know. I can only think of one medication, just off the bat. So I’ll ask them to explain why does he wants to know that. Generally into past medical history by reading systems. They ask general questions that may help us with the diagnosis. | That’s true. Why don’t you examine her? |

The subjects’ criticisms were made in the light of the presented findings, their own differentials, and what they felt the machine’s differential was. Also, the critiques sometimes focussed on the programme environment that they were in. A glimpse at question ((4) How severe is Debra’s headache (on a scale of 0 to 4 with 0 for very mild and 4 for very severe)?) demonstrates that scales in this situation are unusual. However, it is this method of scales which allow such programmes to viably work with fuzziness. For a normal diagnostic interview between a doctor and a patient or between physicians one doctor comments about question (4) as a bizarre way to ask a patient about pain:
How severe is Debra's headache (on a scale of 0 to 4 with 0 for very mild and 4 for very severe)?

So here we have somebody... I don't know why they asked that question unless there is some congenital association between polycystic kidneys and cerebral aneurysms it seems they're both cystic problems, but again the pathogenesis of the cystic kidneys is not the same as that of aneurysms, so I'm going to ask it to explain.

Here the subject (giving the programme credit for knowing something he or she does not) tries to discover the link between polycystic kidney disease and aneurysms.

However, when the interface used "loaded" words like 'necessary', this was perceived by subject 4 as overstating its case. To give an explanation like the one in excerpt 14, is akin to Sherlock Holmes answering to Doctor Watson's question "How did you do it?" with, "Why, it's elementary my dear Watson", and then not telling him why.
The subject then reads on the computer explanation that the reason that the programme “asked” the question was to find out the peripheral white blood cell count.

excerpt 14 continued: So, they want white count, peripheral white cells. They want to know if she has an infection. I agree. That would be useful. O.K. So, we’ll continue.

One concluding consideration, concerning a principle expression of human intelligence, is that subjects also hoped to learn something pedagogically by asking a question. Maybe the machine knew something, which they did not. Staff physicians, in general, tended to develop a working differential around their history more quickly, and thus were less caught with the way the clinical interview went, unless of course if it apparently undermined thing that they needed to round out their differential (figures 6, 7 & 10). Complaints about the order of NEOMYCIN questions concerned the residents, because they worked as if they needed more information to finish their history so that they could go on and develop a solid differential whereas staff physicians were critical in that the interview did not allow them to confirm the differential they were quite certain about. This is a factor of experience.

ANTHROPOMORPHISATION

There is a great deal of interrelation between all of the five categories. One aspect of anthropomorphisation, which is connected with critiquing, is attempting to understand the “reasoning” of the programme, by looking behind the reason for the intentional end of the act as an element of a “person”. As observed in excerpt 9 and 11, the subject assumed what the programmed “physician” was thinking. The subjects were encouraged in this personification by virtue that the programme “explains” itself (which is an attribute of “intelligence”). Self-explanation provides an element of person to person communication.
THE CONSULTING PHYSICIANS

Outside of the experiments, experts were consulted. The 4 specialists in the area of Debra’s illness, who provided the information for the diagnoses and overview of the case of Debra in chapter 4, made their diagnoses with a copy of Appendix A, from which they examined the case. They are described at the beginning of chapter 4. These 3 physicians and one epidemiologist (who was also a registered nurse) acted only as advisors to interpret the case and subject responses. It is interesting to note how they saw the case. They agreed as to the nature of the diagnosis. What was especially impressive about their responses to the case data was that when the case was shown to them, two of them (one neurologist and the infectious disease specialist) felt that NEOMYCIN asked fairly good questions, since it fit well, in general, with their differential diagnosis. Although all of the 4 experts felt that the history and examination questions were incomplete. The other two who were less enamoured with the order of the programme’s questions, thus assumed that this programme’s simulated physician was not the physician who had treated Debra before and therefore the omission of certain paths were a mark of incompetence – that basic questions had failed to be asked. Whereas the other expert physicians assumed that this was indeed the same doctor, and thus, if a question was not entertained, it was because it was not important. Those experts who assumed that this “doctor” already knew the answer also assumed that the previous course of the disease was known by the doctor and that the decisions were “informed” decisions.

THE SUBJECTS

Up until now, with respect to each of the subjects in this study, all of whom are sub-experts, i.e., expert physicians making a diagnosis of a case which exists external to their dominion of specialty, one has seen more than just a shadow of this category of anthropomorphisation. For if the subjects critique the action of the “doctor”, or question how a question should be asked to a “patient” so as to get a usable answer, this
falls into a domain of social context. Taking the reaction of the 4 consulting experts as a guide-line, there are many possibilities for a communication breakdown, since the user can have many attitudes about the perception of the machine—user relationship. (The interface can be perceived as cooperative, or uncooperative, or rather, sensible versus ridiculous).

**The “Physician”**

As for the subjects, the first addressed issue of these results is the perception by the physician subjects of a COMPETENT versus INCOMPETENT "physician".

| excerpt 15 | SUBJECT 4   | The doctor who saw her before this doctor, even ...decided that she needed antibiotics for her bronchitis or for her pharyngitis and a hemophilus influenzae and it went to her ear and it started in her ear |

Subject 4, begins to think about a previous infection which has now spread to the meninges. This is in concordance with the diagnosis of a partially-treated bacterial meningitis, however, at this point, the subject is still looking for confirming information. This subject, later on, comments that the "physician" might not have an otoscope (device for examining the ear), since "it" never looks in the patient’s ear. The experts came to the same conclusion about the partially treated infection, without ever stating a desire to determine the aetiological course of the present disorder, i.e., whether it was an upper respiratory infection or an otitis media, through examination. For the consulting experts the course of the infection that got to this point seemed immaterial; what they indicated that they needed to do at this point was to begin treatment and to determine the type of infectious agent (through laboratory findings).

The difference between how subject 4 and subject 6 treat the interface may very well rest on the fact that subject 4 is very critical of the "physician", whereas subject 6 seems content to try and solve the case from the information that is present since, in fact, part of the work of this physician, as an internist, is to solve "tricky" cases.
Not feeling confident about the competence of the consulting "physician" also accounts for the number of "explanation" requested from the machine by subject 4.

**The "Patient"**

This brings us to the next category, this time with respect to the "patient" Debra and her answers to the clinical questions posed by the programme. The patient was perceived as COOPERATIVE AND KNOWLEDGEABLE versus NON-COOPERATIVE AND IGNORANT with respect to their physical condition. That is, in the discourse between doctor and patient, the patient was seen sometimes as a person whose answers could not be taken at face value. This occurred in places where the doctor had great empathy for the patient. Thus, the subjects did not take the computer interface as a flat field (a "canned" spread sheet) That is, they "brought themselves into" an "lived" in the situation, bringing details from their imagination and reasoning. The display was not perceived as a hard and fast data base with immutable data beyond question. Subject 2 relates, in his or her experience, how questionable patient report data may be:

| Excerpt 16 | Subject 2: But is the stiff-neck true? You have to examine the patient. You can't just ask the patient, "Do you have a stiff neck?" And they say, "Oh, yes I do." What does it mean to them. You can't take that. You have to examine. This is a physical findings. |

Herein lies a case for the subjectivity of certitude with reference to authority. The patient is not perceived as being an authority over his or her condition and thus, the evidence given by the patient verbally, is not taken with certitude. In the case analysis done on subject 4, it is evident, however, that this subject had much empathy for the plight of this patient, and thus did not question the veracity, validity, or reliability of the answers Debra supplies. In the next excerpt the physician indicates that for this finding if a patient generally thinks they have this disorder, they are likely to be right.
Subject 4 doesn't even consider at this point that the 'doctor' would have enough foresight to examine the patient to determine if she had this disorder, a clinical sign which would be noticeable upon routine examination. The biggest unknown for the experimental subjects was whether the information (the findings) were coming from the doctor versus patient.

It is assumed that the data is seen as coming from the patient, but there is no hard reason to rule out that all of the findings could have been derived from the physical examination of the consulting "physician". Again, here the data assuming to be originating from the patient is questioned.

The discourse of diagnostic interview

Thus, the major reason for this phenomenon is that the "situation" mimics, to some extent, a diagnostic interview, and the doctors are thus, in this "mode", so to speak. The discourse between doctor and patient cannot be refined in this interface, and thus the assumption of a such a discourse, could have been making the subjects weary of the robustness of the stated findings. The purpose of a diagnostic interview is to elicit medically reliant information from the patient. The patient does not know exactly what the physician is after, so the physician asks his or her questions in such a way to determine if the information received is of the type sought for. Physical
examination allows the physician to validate immediate medical signs which cannot be feigned, and which are verifiable. However, history-taking makes reference to findings which are proximate, where one has to rely upon the structures interview the patient's memory, understanding, and story-telling ability. The evidence, which is entered into the diagnostic process via this route comes from the ability of the patient, or moreover, the ability of the physician to elicit appropriate responses from the patient in order to make this diagnosis (see page 133 for diagnostic comparisons among physicians).

Doctor-doctor communication differs in that both subjects have a similar knowledge background, thus a conversation to evaluate a patient is directly investigative without having to ask as many question which serve to refine observation related from one physician to another. The person to person interface assumes a great deal of knowledge. However, these refinement questions increase when a subexpert is consulting with a specialist outside of his or her own field, and even more so, when talking with a student. In fact, the student, not being an experienced diagnostician needs to be asked questions which will help in the educative process of becoming a good diagnostician.

Although the computerized physician may not have been heralded by the subjects as a genius, there was only one instance as to whether Debra had more than one headache — (see the section on the CASE STUDY of subject 4 for more detail) where a finding that was derived by NEOMYCIN was questioned, even though the reasoning was often criticized.

ATTRIBUTES OF PERSONIFICATION

The results indicate that part of dealing with the interface depends upon WHAT is being asked told or commanded by the linguistic interchange. Some of the things which affected the doctors were in part turned towards what the different "people" were doing during the experiment, and what was expected of them; e.g.,
The force of the question: "How would you characterize the onset of Debra's headache?" seems like commands. It is, of course, a request, which is not, however, aimed at the subjects in this study.

Therefore, as hypothesized, in addition to making considerations about themselves and the presented medical findings they also made assessments about homuncult which emerged from the programme. They reacted to the sentence forms, with respect to what they assumed about the characters, and to what situation was. They perceived Debra, the patient, the consulting physician, and the physician who saw or treated Debra before, who might be the same one who is looking at her now. The issues which they have addressed in their protocols have denoted a concern for WHO is asking or telling or commanding.

This is also true for WHEN (or rather at what point in the diagnosis) is it being asked told or commanded and HOW is it being asked told of commanded. This however rested more with their domain-specific reasoning. Whereas, WHY is it being asked told or commanded leads them to think about the intent of personages, and this, affects their reasoning, their explanation as well as their critiquing, especially in terms of the computer question and explanations.

Since the programme can ask a question, and can explain its question. The subjects are in a completely interactive interface with the programmes rationale for its behaviour. Because of the fact that there are clinical personages in the programme interface, the subjects paint the characters in the programme with intentionality. This
is marked by their conversational conventions as well as expectation that the subjects have about physicians and patients and what they do in these situations.

It is important to note that although NEOMYCIN can interpret its problem solving inquiry, for those physicians who asked for an explanation of a finding expected reference or pedagogical information. Although it could not directly settle points of confusion about findings, misunderstanding in this regard could be alleviated indirectly through explanation.

Because the subjects treat the interface with similar expectations, which they do of other interpersonal communication, the results reveal instances of indices of clausal ellipses which can appropriately describe the of their verbal data.

**A TEXTUAL ANALYSIS OF SUBJECT’S REJOINDERS**

As we have seen, of any item of information which the physicians encountered on the computer screen, they either accepted that information on face value, rejected it or were confused by it. The possibilities of informational acceptance, rejection, or confusion rest with the salience of the interface as well as with the subjects mental models. Taking note of subjects’ reactions to NEOMYCIN presentations proved a authentic way to keep track about the clinical reasoning as well as pragmatics. The subjects’ interaction with the ‘explain facility’ is like asking, "What do you know that we do not?" It is this filling in the blanks that subjects have to do if information is not supplied to them. Mainly it was the subjects disagreement, surprise, or confusion with a piece of textual information which lead to their use of the explanation facility. These objections produced the most conversational-type responses.

**Instances of Indirect Rejoinders**
In the example below, one can see an instance of SUBJECT EXPLANATION which are comments in the Halliday & Hasan (1976) clausal ellipses. These commentary made by a doctor in which he or she implicitly accepts the presumed clinical intent of the COMPUTER QUESTION, "Does Debra have an abnormal funduscopic exam?" as being a move by the "physician" to investigate the possibility of increased intracranial pressure. This questions follows the finding of an abrupt headache, which coupled with a severe headache implies a cerebral haemorrhage. Thus, the order of the textual sequence, and the present differential of subject #3 - a medical resident, lead to following response:

**excerpt 23** [SUBJECT 3 I was gonna, I was starting to be very interested in the physical examination because... it brings it back to a more serious level again]

What this means is that subject 3 thinks the disease is more serious now, in light of the new evidence, and that the "physician" thinks so well- i.e., that the questioning is not random or directionless, moreover the subject states:

**excerpt 23 continued** [SUBJECT3 And he asked about physical examination... abnormal funduscopic exam because he's thinking about increased intracranial pressure].

An abnormal funduscopic examination would be an indication of an intracranial haemorrhage).

The phrase uttered by subject 4— it is a good question, but I would not ask it now— is an example of COMMENTARY indicating the subjects attitude about the programme's question. Much like the above example of partially rejecting a question, acceptance of a question does not so much address the content of the question, as its utility and the assumed intention of the programme 'physician'. It is a weighted rejection of the appropriateness of the question #12) Does Debra have a history of polycystic kidney disease?
Looking again at the context from excerpt 17 shown below, when the physician tries to tap into the programme's reason for asking this question at this point in time, it is assumed by the physician that the patient did not give answer to the programme physician's question. The subject projects onto the 'physician' as being incompetent and assumes that the programme "physician" is not satisfied that the information coming from the 'patient' as being complete. Thus, subject 4 is making a clinical judgement, or rather giving clinical advice about instances where patients give certain types information. This is much as subject #2 did in reference to the headache (see excerpt 16), except, for subject 4 below, for this symptom is on the side of the patient's opinions which indicates a real-life social linguistic situation has emerged:

<table>
<thead>
<tr>
<th>So, Let's see what they think. Why they think it's a good question. Why are we asking whether Debra has a history of Polycystic kidney disease? We are trying to determine whether Debra has Polycystic kidney disease</th>
<th>Meaning that patients don't know when they have it. But, I bet when they they say, &quot;Yes,&quot; they are right more often than if they say, &quot;No.&quot; They [the computer physician] may not know.</th>
</tr>
</thead>
<tbody>
<tr>
<td>NO</td>
<td>O. K. So let's...We'll go on. She doesn't</td>
</tr>
</tbody>
</table>

This statement seems to involve more that just the nature of this medical condition The very next question that the programme presents also implied that Debra's "No" might not be so, that she might not know that she has this disorder. The question was 13) Does Debra have a family history of polycystic kidney disease?

| excerpt 24 Subject #4 Does Debra have a family history of polycystic kidney disease? I don't know if that's familial. I think it might be. I'll see what they say...We are trying to decide whether Debra has polycystic kidney disease. A family history of polycystic kidney disease is weakly associated with polycystic kidney disease. |
| --- | --- |
| So, the answer's, "No" (in reference to the previous NEOMYCIN question as to whether Debra has polycystic kidney disease PCKD). It might mean she doesn't know and it might mean the answers no, but if the answers yes she probably knows. **NO** |
| She doesn't have a history. |
These experimental findings relate to the Gricean Maxim of relation which states that one assumes that an utterance is relevant to the linguistic and extra-linguistic context (Grice, 1975). The immediately preceding linguistic content being the most candidate. Moreover, it follows that a speaker would also assume the same about an interlocuter. The nature of interfacing with a machine implies that the machine makes a response implying an adequate interpretation of the human being's prior actions' significance (Suchman, 1987). The assumption is that everybody is working in the same context, NEOMYCIN, the "patient," the "doctor" and the subject. As we have seen, this also holds true for interchanges with the programme's explanation module.

With respect to excerpt 24, and the Gricean maxims of quantity (informativity): one may assume that if an alleged sequence of utterances between the "physician" and the "patient" (which we do not see transcribed on the computer screen) were immaterial, asking about family history of PCKD should have come before asking about it in an individual as one subject noted. When this was not assumed, the second question appeared senseless to the subject. In other words if the doctor wanted to know if Debra had hypertension (which would be far more related to her current health and predicament than PCKD), then he should have simply have taken her blood pressure, than asking if she had this condition which would give her high blood pressure. Moreover, if one is curious about PCKD, one would ask about it in the family history first. To do it backwards leads one to think that there is more here than meets the eye.

In terms of these results, a switch occurs in the above instance where a physician makes note that if a patient told the physician this piece of information it would most likely be true. Whereas the opposite is strikingly the case here:
She doesn't have any visual problems. That's interesting. I would have thought that...if he asks her, in that way, "Do you have any visual problems?", and she said, "No...I would... — just, "No. "...If that's her only answer, with no elaboration, I would have interpreted it in a different way. I would have used that more to decide whether or not I though it was a migraine, or whether or not it had any sort of seizure-like activity (or an aura of a migraine). And without that, then I would have taken it... If its just nil, I would still have interpreted it as an aura and decided that she didn't have one, and therefore that it was less likely to be migraine, and I would, I would be more likely to think... she...she still be considering meningitis in spite of no fever and no seizure.

CONFUSION

In the previous example of abnormal mental status (excerpt 18), there was an instance of semantic confusion. Confusion was not only in reference to not understanding what is being said, but about what was being done. In the case of subject 6 mistaking an abnormal mental status, for psychosis, the explanations cleared up the misunderstanding. Subject 4, on the other hand, ran into a rare nuance of NEOMYCIN showing its pedagogic and programming end. The last question is a task consideration as to the subject's own role in the interface.

So, why is he asking that? Um, Severity...We are trying to learn more about the headache. Headache severity is a characteristic feature of a headache. Let's see if they'll. We are trying to give thorough consideration...headache is a new finding that needs to be clarified first. I don't know what that means. What else? I'll ask it to explain one more time. The user must state the problem. They want me to punch in the reason why I'm having trouble with it? Well I don't know what they're doing. Am I the doctor?

AN EXAMPLE OF DISCLAIMER-REJECT

What follows is the only instance of the object of a question made by the programme being questioned in essence, it is further analyses in the first case study.
A Disclaimer is side stepping a question by disputing it's relevance. This category is close to what is demonstrated in (excerpt 26) is: Does Debra's headaches have ... She doesn't have headaches.... The subjects saw evidence for only one headache episode of note and thus rejected the underlying premise of multiple headaches in the programme's question. But this example does not fit the Disclaimer category exactly, since the question is not merely side-stepped, it is directly challenged as inappropriate or mistaken in its presupposition. The statement, serves as an answer to tell the programme-questioner that the headache is singular. The last commentary made by the subject about the headache not hurting does not appear to make sense since a subarachnoid haemorrhage should produce an excruciating headache, but the reference is most probably related to exacerbating that headache, or to the type of headache which could be exacerbated or have precipitating factors.

To sum up, when the subjects tried to guess the medications; this was a direct response to a question, as well consent to a command. There were no instances of supplementary responses to a question, except if one would want characterize all of the commentary (in the form of SUBJECT CRITIQUES and SUBJECT EXPLANATION) as the principle indirect responses. This would not be the way Halliday and Hasan (1986) would use these terms as these were not usually responses to questions, but to statements. It seems the best way to treat subject responses is as though they are responses to questions. This serves our purposes of differentiating responses in light of the fact that the subjects are being probed to think aloud and therefore are answering a
directive if not a perpetual set of questions, and commands (although they were not being reminded of the probes unless they stopped thinking aloud). Although this seems trivial, the fact is that the subjects were never told, in any instance to answer NEOMYCIN's questions, which were never directed towards the subject in any direct way.

**Responses to Commands and Statements**

As above the subjects assented to all of the direct and implied commands in the experiment. Virtually all findings were taken as valid if coming from the "physician", but the subjects were wary of those they felt originated from the patient. It is difficult to say, definitively, why the subjects felt that a certain characters would addressing certain things at certain times, and not others.

The subject never contradicted a command, nor a (non-interrogatory) statement made by the physician, but the same was not true, as demonstrated above, for patient statements.

The only quasi-statement of a "physician" that was contradicted was the implication of a question which presupposed Debra having more than one headache.

**Question Rejoinder**

There are two types of question REJOINDERS. One that asks a question, the other which makes a supplementary statement. When the programme was not doing what was felt by the subjects as essential physical examination to find the origin of Debra infection, subject 4 said out: "Why don't you look in her ear? and then later said, "He doesn't have an otoscope — this doctor, which are respective examples of the two types of question rejoinder.

**CONSENT**

By the fact that the subjects did all that they were told and did all that they thought the programme asked of them it can be assumed that they consented to the commands of the probes. NEOMYCIN presented only one string which was in the form
of a command (the others were declarative statements): **Please elaborate.** This was with respect to the medication which Debra may or may not have been taking, surprisingly the subjects all did their best to elaborate; trying to think of all the medications she could be on that could have produced those symptoms, or ones she could have taken to relieve her symptoms. They may have interpreted — please elaborate — as one of the experimental probes. Thus, the subjects commented there for there were no instances of refusal to perform a command.
CASE STUDIES

CASE SCHEMATA OF EXPERIENCED PHYSICIAN

One subject was chosen for a more in depth analysis of the medical reasoning, since it was the subject who elicited the most striking examples of interactive critiquing and anthropomorphisation. Upon a closer examination of subject #4, one can observe the medical reasoning that was taken at each step in the history and the physical examination of Debra. The schemata (figures 5a through 5f) show the clinical consideration made by this physician while the history of Debra was unfolding. These diagrams follow a convention.

The summarized content within the square boxes were those which originated from the compute screen. These boxes contain observations. The term OBSERVATIONS refers to the data which was presented textually on the computer screen, as the subject worked with NEOMYCIN. These include findings as well as the medical questions produced by the programme. The thicker, darker square boxes represent the findings of the patient, Debra, whereas the lighter, thinner boxes contain questions posed by NEOMYCIN. The rounded boxes contain information coming from the subject’s verbal protocol. One such piece of information from the subject is the DIFFERENTIAL which refers to the hypotheses which this doctor stated he or she was considering at a point in time. The box labelled OTHER CONSIDERATIONS refers to that subject’s relating of his or her observations to the entertained differentials. The arrows show cohesive reference from the observation to the reasoning consideration. One sees that as early on as with the patient description (figure 5a) the fact that NEOMYCIN has given the patient a number (in its data base) lead this physician to suspect that Debra may have been previously hospitalized — e.g., Does that means she’s been in before? — and that this number was thought to be her hospital number.
Does that mean that she has been in before?

Does that mean that she has been in before?

Does that mean that she has been in before?

Does that mean that she has been in before?

Does that mean that she has been in before?

Does that mean that she has been in before?

Does that mean that she has been in before?

Does that mean that she has been in before?

Does that mean that she has been in before?

Does that mean that she has been in before?

Does that mean that she has been in before?

Does that mean that she has been in before?

Does that mean that she has been in before?

Does that mean that she has been in before?

Does that mean that she has been in before?

Does that mean that she has been in before?
This consideration would make a difference in the instantiation of the subject's initial differential diagnosis of Debra's condition; i.e., whether Debra has a chronic treated problem versus an acute untreated problem.

If Debra was just coming into the emergency ward with her symptoms, this physician, like all of the physicians in this study, felt that the history-taking and the physical exam were neither complete, nor in a chronological order, which they would have done or recommended.

The presenting headache, stiff neck, nausea and vomiting elicited a differential which included the consideration of an infectious process (indicated when this subject inquired if Debra had been feeling feverish, or feeling like she had a flu). This remark was probably made in light of the meningeal irritation, and the possibility of severe hypertension. The thing which is quite impressive, in the doctor's reasoning, is that although pregnancy is never explicitly, not implicitly mentioned by NEOMYCIN, with respect to Debra, and that the subject could clearly not know if Debra was pregnant, this physician brought the consideration of a problem in pregnancy into the differential and retained it there throughout the case. Thus, with regards to the issue of competence versus incompetence, with respect to the computerized 'physician', it would be reasonable to expect that the subject felt that a competent physician would have pointed out the fact that a patient was pregnant. This, hence indicated a lack of trust for the information given by the "physician". There are no circumstances (even in cases of perfect health) were this condition is not relevant (for we are dealing now with two patients not one). The (hypertensive pregnancy) abnormality which subject 4 considered is termed preeclampsia - and is treated with, among other things, a magnesium sulfate drip- (see figure 5b). This abnormal condition, moreover, occurs in very late pregnancy, where someone would be obviously pregnant. The only other assumption would be that the subject's behaviour is a frustration response at not being able to ask this obvious question in the experimental trial.
**OBSERVATION**

**DIFFERENTIAL**

- Headache Severity?
  - Severe Headache (on a scale from 0 to 4)
  - Migraine

**OTHER CONSIDERATIONS**

- She (the patient) thinks she's having a stroke or something.
- Headache is severe: for a 6 hour headache no one would come in unless they were really sick, but people with migraine don't come in for a couple of days.
- Headache must be severe, or would not have come in.

**TREATMENT**

- We are going to call an ambulance, and she needs to be on a magnesium sulfate drip.

**Figure 6b** Schematic of Clinical Reasoning in Gerontologist Subject #4 (continued)
This latter hypothesis is unlikely since the pregnancy hypothesis occurred before the doctor started physically operating the machine (see figure 5a).

The severity of the headache lead to the addition of migraine to the subject's differential. Although this physician felt that the question about headache duration, and severity lent nothing to the clinical picture, the question about severity, at least, stimulated the physician to talk about the consideration of a possible migraine, which had not been spoken about before that.

This physician's protocol for the question of headache severity (figure 5b) denoted urgency (e.g., **No one would come in unless they were really sick**). This physician empathized very much with the "patient's" plight — (anthropomorphisation). So, with respect to the patient being knowledgeable versus ignorant issue, it is assumed by this subject that the patient was not malingering. Therefore, the patient-stated level of the severity of the headache was informed and accurate, and not just a bad tension headache, and that the patient was really sick and needed to be treated right then, as she was in agony. (This was in contrasted with subject 2 who considered that the patient may be malingering — see excerpts 15 &16). Moreover, subject #4 was the only physician who offered a treatment. However, this treatment was made, not only with respect to the data presented in the case, but furthermore, in what one could term a possible world. or situation. The assumption was that if Debra is pregnant, and had these symptoms, then she should be treated thus. The physician did not necessarily believe that Debra was pregnant, however. the physician just didn't know. Again, there is no reason to assume that the patient was pregnant, if the 'physician' did not reveal it, since pregnancy is a complicating factor of all levels of health, both with respect to the mother and to the child, at all stages of pregnancy.
OBSERVATION

Differential

No Recent Seizures:
- Migraine
  - Very bad headache which won't go away, with a lot of side-effects

No Fever:
- Still could have Meningitis
  - We are back to migraine still possibly a viral meningitis; though less likely
- She (the physician) could still be considering meningitis in spite of no Fever and no seizure:
  - Aura of Migraine

Visual Problems?
- Migraine less likely.
  - Meningitis
  - She is not pregnant and had seizures. I think Meningitis, but it is less likely; she does not have a fever.

No Visual Problems:
- Migraine less likely.
  - Meningitis
  - He's (the physician is) probably asking that because he's considering migraines
  - No aura
  - He's wondering if it's migraine as a symptom of migraine or could the visual problems be an exhibit of seizures

Headache Onset:
- Abrupt?
  - Sudden?
  - or Gradual?
  - Aneurysms
    - (Ruptured — Hypertension)
    - So now I'm thinking aneurysm.

OTHER CONSIDERATIONS

Figure 5c Schemata of Clinical Reasoning in Gerontologist
Subject #4 (continued).
Like all of the subjects, this doctor considered meningitis, intracerebral bleeding, and migraine as the differential. The fact that Debra did not have a fever lead this subject to reshuffle the priority of the items in the differential (see figure 5c). Migraine, is less likely, as Debra does not manifest any visual problems. With NEOMYCIN's presentation of the next question about visual problems the subject made an assumption about what the computerized "physician" was "thinking" when he or she asks that question. This is an example of critiquing and personification. The personification is that the thoughts of the 'physician' are being considered. The criticisms of the thought content constitutes critiquing. The physician had built up scenarios in the consideration of items in the differential; one being that the patient had a migraine, thus relegating the other symptoms to the role of side effects of that main disorder, migraine. Hence, in such an instance, one can see that once one hypothesis gets more consideration, the other pieces of evidence may take on different roles in each hypothesis. For one hypothesis a symptom may be primary, for others it may be secondary.

The discovery of the abruptness of the headache brought forth a hypothesis of intracranial bleeding (aneurysms) in this subject and for all the subjects. For some of the doctors the question of hospitalization made them wonder if she was previously treated, but others only referred to the usefulness of the question itself. Figure 5d shows a continuation of the physician's hypertension theory which had remained in the differential since the beginning. There was, now, a consideration of a host of vascular problems, including migraine. Although, in the beginning, hypertension appeared tied in with pregnancy, the consideration of hypertension was also considered without the patient being pregnant. The possibility of high blood pressure bursting an aneurysm was thought about. The interesting thing here was that the subject asked the programme to explain why it was looking for an abnormal funduscopic examination result, and why it was asking as to whether the headache had precipitating or
aggravating factors. NEOMYCIN's considerations, made ostensible through the explanation facility, are italicized in the OBSERVATIONS boxes. The subject agreed with the programme that the ophthalmological (eye) exam could be an indicator of a retinal haemorrhage. What had occurring here was not that the subject thought that looking into the patients eyes was a brain storm. (Recall that this is the same physician who thinks that the NEOMYCIN doctor should have looked in Debra's ear). It was felt that an eye examination here would have been appropriate. Such an examination is a typical and routine procedure performed in all physical examinations, thus, the subject's request for an "explanation" was more due to the fact that the subject had made a presupposition as to the illocutionary force of NEOMYCIN's question. The doctor then assumed that the programme was anticipating that Debra might indeed have an abnormal exam. So when the subject selected "explain"; that action, could be interpreted as the lexical string, "Why do you suspect that Debra has an abnormal ophthalmologic examination?", rather than, "Why are you asking for this test now".

However, this physician was sensitive enough to notice that the interview, up until that point, made reference only to the one presenting headache. So that when the programme asks about whether Debra has headaches with precipitating or aggravating factors, the subject asked again for an explanation. In response to the subject's request for an explanation, NEOMYCIN indicated that it was considering a subarachnoid haemorrhage. If the presenting headache was indeed the result of a possible subarachnoid haemorrhage, then this whole situation semantically entailed that Debra could only have had the one headache of interest – i.e., the presenting one, which is being considered in the possible world where the above would hold true.
OBSERVATION

- Funduscopic exam
  - Abnormal?
  - Explanation: Retinal haemorrhage

DIFFERENTIAL

- Hypertension
- Migraine

OTHER CONSIDERATIONS

- She probably has high blood pressure, that's why she ruptured the aneurysm.
- Would be a sign of high blood pressure.
- Less towards Blood pressure
- Very sudden headache does fit with migraine
- Toxicity

- So Debra's headache(s) have: O.K., we have not heard that she had headache(s) of any kind before, and this is assuming that she had them before.
- Maybe gets headaches from asthma medication
- She could have a sudden headache caused by subarachnoid haemorrhage, that is true
- Assuming that she had headache(s) and she has subarachnoid haemorrhage, which I think doesn’t fit... if we are postulating that she has abrupt onset of subarachnoid haemorrhage, then she only had one headache ever. She doesn’t have headache(s) that have precipitating factors

Headaches with precipitating or aggravating factors?

- Explanation: Subarachnoid Haemorrhage

Headache with no precipitating or aggravating factors?

- Abnormal mental status?

Figure 5d. Schemata of Clinical Reasoning in Gerontologist Subject #4 (continued).
However, the entailment becomes merely a presupposition which rests upon the assumption that the subarachnoid bleed referred to the patient’s true condition in the real world. There is no reason not to suppose that the patient could indeed have had a series of headaches; this last one being the one that sent her to the doctor. However, if this is true, it is also true that virtually all people do not have a series of subarachnoid haemorrhage related headaches, and live to talk about it — not to mention be able to make it to a doctor. The medical evidence of people having a ‘leak’ headache (which is nowhere as severe) which “warned” of an ensuing subarachnoid haemorrhage is typically postdictive, and would only be a totally academic consideration here. Thus, the subjects’ response was COMMENTARY; a reaction to the embedded phrase Debra’s headaches in the explanation of the programme. It is an instance of REASON with respect to a subject EXPLANATION. That is the subject stopped to consider his or her own picture of the case with that of the programme. In spite of this, it is noteworthy that question 3) “How long has Debra had this kind of headache?” did not lead the subject to think about other kinds of headaches.

Now question (3) “How long has Debra had this kind of headache?” does not necessarily have to be taken to mean (3i) How long has Debra had this headache?, nor does either (3) or (3i) have complete congruence with (3iii) How long has Debra had the headache?. The string 3ii is neither the same as (3iii) How long has Debra had a headache? Yet, in this context, the physicians dealt principally with the headache of the presenting complaint, even though some of the other physicians would have liked to know more about Debra’s headache history.

The question of WHO is giving information about the headache; whether it be another physician or Debra herself may be denoted as (3iv). How long has Debra reported that she had this kind of headache? The answer to (3) that is given by NEOMYCIN is an appropriate answer and in an expected format: of a WH-QUESTION (3a) 6 hours = Debra has had this kind of headache for six hours. Although it is also a
viable answer for (3i) through (3iv); each interpreted question can elicit different sorts of commentary. For (3) — What kind of headache is this? — can inspire: "Does she have other types?", and so on. For (3i) and (3ii) there is no ambiguity about which headache we are referring to, but due to the circumstances neither is there this fuzziness for (3iii). However, (3ii) can be responded to as though it were "How long has it been since Debra has had a headache?" which weakly connotes "What circumstances preceded the headache?"; possible answer: "Ever since she ate those melons". Which would make it more like question (10): Do Debra's headaches have precipitating or aggravating factors?

Figure 5e represents a suspension of judgement due to the subject's response to the programme asking a set of questions which the subject did not feel were very relevant to the case, at that time, except that polycystic kidney disease is associated with hypertension. The physician felt that Debra might have been ignorant of having a polycystic disease condition (see excerpt 17) and thus, she could not answer the programme's 'physicians' question. In Figure 5f NEOMYCIN pursued which medication the patient might have been been taking, the reason being the 'physician' does not know what they were. However, the results ended up determining what medications Debra is one to treat her condition, specifically. The 'physician' pursued this avenue as well, but the subject's intention is to determine what side effects of the medication could produce the symptomatology. Again, in Figure 5f it is assumed that if Debra was medicated, but not hospitalized then she has been treated for a mild infection. Consider that the consulting experts and all of the subjects considered bleeding on account of the abruptness, and the partial treatment by reason of of the medication, indicating that minimal information was used to trigger a schema. Although this schema is not true out of necessity, but a minor infection more likely in a case of these circumstances.
One can see that the subject did not have a tremendous amount of information on the patient, and therefore had to rely more upon discerning judgment: to make selective inferences and assumptions. Indeed, to a novice, leaps of faith. Much of the information on the patient were negative findings rather than affirming information and the source of patient information was not clear. Did it come from the patient, or from the doctors? Theses factors without adversely hampering the subjects medical reasoning, produced so many assumptions and expectations from the physician that these things, in and of themselves, are worthy of note. Again, in Figure 5f it is assumed that if Debra was medicated, but not hospitalized then she has been treated for a mild infection. Consider that the consulting experts and all of the subjects considered bleeding on account of the abruptness, and the partial treatment by reason of of the medication, indicating that minimal information was used to trigger a schema. Although this schema is not true out of necessity, but a minor infection more likely in a case of these circumstances.

One can see that the subject did not have a tremendous amount of information on the patient, and therefore had to rely more upon discerning judgment: to make selective inferences and assumptions. Indeed, to a novice, leaps of faith. Much of the information on the patient were negative findings rather than affirming information and the source of patient information was not clear. Did it come from the patient, or from the doctors? Theses factors without adversely hampering the subjects medical reasoning, produced so many assumptions and expectations from the physician that these things, in and of themselves, are worthy of note. It is difficult to say why this one physician trusted all the information as coming from the patient, whereas others did not, or why the doctors assumed that some things were the product of the considerations of the computer 'physician', whereas some others saw them as being supplied by the patient.
Observation

History of polycystic kidney disease?

Differential

+Hypertension

+Looking for cause of high blood pressure

+She probably doesn’t know

Other Considerations

Family history of polycystic kidney disease?

No family history of polycystic kidney disease?

+I don’t know if that’s familial.

Figure 5e  Schema of Clinical Reasoning in Gerontologist Subject #4 (continued)
OBSERVATION

Taken Medications Recently?

Yes, Took medication recently

Elaborate on medications?

Patient took:
ANTIMICROBIALS>
ANALGESICS>

Differential

Toxicity

Medication for asthma caused headaches in young people: she's not so young — 39. Street drugs

Patient took:
Partially-Treated Meningitis

Viral Infection

Differential

OTHER

CONSIDERATIONS

 Been hospitalized Recently?

Has not been hospitalized recently>

Since she has not been in hospital this would be a minor infection

Probably for meningitis infection, but perhaps she has a viral infection. The doctor, who saw her before this doctor, decided that she needed antibiotics for her bronchitis or her pharyngitis and a hemophillus influenza and it went to her ear and now it is in her meninges a bronchitis and it was going to become a meningitis

Figure 5f. Schemata of Clinical Reasoning of Gerontologist
Subject #4
One thing that can be said is that these clinical considerations depended upon the focus of the physician, not only with respect to the differential, but in what that individual found salient in the patient information. Whether the focus be clinical practise or pedagogy an analysis of bias and prejudice could start here to see if individual difference is the source of particular attitudes. Some individual may be more analytical, and some more intuitive, for instance. In addition, this way of using our intellect and the knowledge which we have must be represented somehow in the human being who has this understanding, and this shows through in the behaviour in the microworld. Could it be that affective variables as in (Dreyfus, 1981) need also to be considered?

This physician, despite seemingly being fixated on pregnancy, and the computer explanation facility, has the same differential as the other subexperts, and even considers the actual diagnosis of partially-treated meningitis. However, like all the subexperts, the physician is not certain of this diagnosis.

COMPARISON OF NETWORK REPRESENTATIONS

If we take a look at a representational network (figure 6) of Subject 4’s reasoning up until the time he or she discovered that Debra does not have a fever, one can see that the items in OTHER CONSIDERATION are making reference mainly to the items in the DIFFERENTIAL. These are hypertension (with pregnancy), infection, and an infectious process (with meningeal irritation). Debra is more likely to be hypertensive because she is black, but the physician never mentioned this in the considerations. It is from this picture of hypertensive pregnancy causing probable seizures (considered when the question was asked about seizures) where subject #4 considered a treatment (see figure 5b).
Observations

1. 39 Year Old
   - Lady
   - Negro

2. 6 Hour History Of Headache
   - Stiff Neck Associated
   - With...
   - Nausea
   - And
   - Vomiting

3. Headache
   Duration: 6 Hrs.

4. Headache
   Severity: Very

5a. Fever?
5b. No Fever

6. No Recent Seizures

7. No Visual Problems?
   - photophobia

Differential

- Infectious process
- Meningeal symptomology
- If pregnant → Severe Hypertension
- Severe headache.

Other

Considerations

- Migraine
- but people with migraine don't
  come in for a couple of days.

Treatment

- Meningitis less likely
- by now I would know if
  she was pregnant or not
  so I could drop that hypothesis
- Infectious process
- If she is not pregnant and had seizures
  I think of meningitis, but it is less likely
  since she does not have a fever.
- Viral Meningitis
- If she is pregnant and had seizures
- we are going to call an
  ambulance and she
  needs to be on a
  magnesium sulphate
  drip.

Migraine
- menigitis
- no aura

FIGURE 6. NETWORK SCHEMATA OF MEDICAL REASONING SUBJECT #4 - GERONTOLOGIST.
OBSERVATIONS

2) 6 Hour History of Headache
- Stiff Neck
- Nausea
- Vomiting

3) Headache Duration: 6 Hrs

4) Headache Severity: Very Severe

5) No fever

6) No Recent Seizures

7) Visual Problems?

DIFFERENTIAL

Intracranial Problem
- Infection
- Meningitis Encephalitis
- Space-Occupying Lesion

OTHER CONSIDERATIONS

Vascular Problem
- Subarachnoid Haemorrhage caused by
- AVM's & Aneurysms

FIGURE 7. NETWORK SCHEMATA OF MEDICAL REASONING SUBJECT #1. Gerontologist
Similarly, another experience physician (subject 1—see figure 7) related the other considerations to the differential. The differential of this physician included the possibility of a space occupying lesion (which is any abnormality in the intracranial area which takes up room) and infection; both of which are very vague global categories. This physician considers bleeding more prominently that subject 4 does. However, again as subject 4, the new information subject 1 attains is related to the current differential which is being considered. The programme's question about the possibility of a seizure is considered only in connection with its association to the physician's differential (seizures being associated with aneurysm being in themselves the cause of a subarachnoid haemorrhage).

These more experienced physicians performed slightly differently from the least experienced resident—subject 2. In striking contrast (see figure 8) the medical resident (subject 2) makes many more considerations referring back to the presenting symptoms per se rather than to the initial differential. When one looks at this physician's reasoning via the medium of a graphic network, with respect to the last two figures (figures 6 & 7) of the more medically experienced physicians, it is notable that there emerged two types of patterns for trying to solve the case. Figure 8 shows that subject # 2 refers most of the medical considerations to the initial symptoms of headache, nausea and vomiting. Looking again at subject # 1 (see figure 7), it can be seen that the consideration herein revolved mostly about the hypotheses of lesions infections, and bleeds that have been made about the illness in the differential. The process of relating things back to the initial symptoms was characteristic only of the early residents in this study—subject 2, but not the more advanced residents subject 3 and 7 (see figures 9 & 11, respectively). However, they were more sensitive than the staff physicians about the way NEOMYCIN was diagnosing. The hypotheses-oriented diagnostic strategy was employed by the experienced physicians—subjects 1, 6 and 4 (see figures 7, 10 & 6, respectively).
Observations

1. Year

2. History

3. Headache

Other Considerations

- Hypertension
- Meningitis
- Encephalitis
- Headache
- Duration
- Intraocular bleed
- Trauma
- Meningitis
- Meningocele

Differential

FIGURE 8. NETWORK SCHEMA OF MEDICAL REASONING SUBJECT #2 - RESIDENT.
Observations

1) 39 Years - Lady - Negro
2) 6 Hour Headache
   - Stiff Neck
   - Nausea
   - Vomiting
3) Headache Duration: 6 hours
4) Headache Severity: Very
5) No fever
6) No Recent Seizures
7) No Visual Problems

Differential

- Meningitis
- Less likely stroke
- Migraine
- Temporal Arthritis

Other Considerations

- Redundant question
- Maybe headache is secondary
- Maybe headache is primary and nausea and vomiting are secondary
- He's thinking infection.
- I would ask about photophobia, exposure to infection, sore throat
- Meningitis not secondary

FIGURE 9. NETWORKSCHEMATA OF MEDICAL REASONING SUBJECT #3 - RESIDENT.
Observations

1) <39 Years> <Lady> <Negro>
2) <6 Hour Headache>
   • Stiff Neck
   • Nausea
   • Vomiting

4) <Headache Severity: Very>

5) <No fever>

6) <No Recent Seizures>
   • Cerebrum intact

7) <No Visual Problems?>
   • Less likely vascular, neoplastic, or severe infection

Differential

Infection
- Meningitis

Hypertension
- Intracranial Pressure
- Stroke

Generating history and physical examination questions

Other Considerations

- Headaches questions
- Question about the stiff neck
- Fever
- Exposure to infection
- Level of consciousness
- Hereditary disorder
- Focal signs

Blood pressure

Laboratory test: Lumbar puncture

Seizure

And that would give me the diagnosis

FIGURE 10. NETWORKSCHEMATA OF MEDICAL REASONING
SUBJECT #6 -INTERNIST.
OBSERVATIONS

1) <39 Years>
   - Negro Lady
2) <6 Hour History Of Headache>
   - Stiff Neck - Nausea - Vomiting
3) Headache Duration: 6 hours
4) Headache Severity: Very
5) No fever
6) No Recent Seizures
   - meningitis
7) Visual Problems?
   - no Visual Problems?
   - photophobia

DIFFERENTIAL

- Migraine bleeding into the cerebrospinal fluid
- Meningitis
- Subarachnoid hemorrhage
- Vascular
- Tumour mass
- Migraine
- Bleeding
- Intracranial mass

OTHER CONSIDERATIONS

- Good to ask again, often times the patient will remember the headache started before
  - if less severe migraine
  - if very severe new disorder
- Ask about seizures,
- 3rd nerve palsy
  - double vision

Assumes NEOMYCIN thinks she has severe meningitis because asked if she had photophobia, but she has no seizure/rash
- Condition not severe

FIGURE 11. NETWORK OF MEDICAL REASONING SUBJECT #7 RESIDENT.
The different experienced physicians may have not been utilizing different techniques, but the difference may have denoted the same process, but with the more experienced physicians having a greater a underlying capacity to use familiar information (chunks). It is suspected that the more experienced physicians have a much more highly compiled history-taking and physical examination stratagem which is not directly demonstrable in the physicians protocols, but from their ability to look at symptoms, signs, disease conditions and hypotheses simultaneously.

The consulting specialist, who were experts in this field, were able to catch on to salient information, selectively focusing on relevant information, even better than these experienced subjects. The non-experts must, on the other hand, explore multiple pathways, as subject 2 did. Patel, Kaufman, & Evans. (1988) show that, in live doctor/patient interactions, medical residents will ask more irrelevant questions when taking clinical history from a live patient. In this study, the experimenter, (Kaufman, 1987), had physicians problem-solve in a live doctor/patient interview. Below is a transcript of a resident at the beginning of a clinical interview with a live patient, which is indicative of novice interview behaviour:

Doctor:
What's the problem?
Do you have muscle problems in your legs? What do you have muscle weakness, or?
No pain at all. how long have you had the weakness?
Three months is it there all the time, or does it come and go?
OK. Do you find let's say... Do you have the weakness when you wake up in the morning?
Do you have to do a certain amount of exercise before you get weak?

Patient:
I have muscle problems? Weakness, no pain.

Three months
It's coming and going. It's happened twice.
No.

Ah... No.

from Kaufman (1987)

The question which the resident asked are focussed on trying to explain a single finding – muscle weakness (Patel, Kaufman, Evans, 19889). The resident does not attempt to elaborate upon the presenting complaint in order to investigate other
possible symptoms. Since this, as well as the other residents in Kaufman's (1987) study pursued multiple solution pathways, they asked more questions than the expert physician. The near expert (sub-expert), if he or she does not figure out the case immediately will have to rule some things out and build a case, in terms of the existing differential. These more experienced physicians, in the present study, who can be regarded as subexperts, or low-domain knowledge experts (Joseph and Patel, 1990) do just this. Novices do not seem to recognize the significance of latter evidence with respect to their differential, and thus change direction. Experts have rich models of their domain (which may be incomplete or inaccurate) (Johnson-Laird, 1983). Thus, with the exception of subject 2, all of the physicians, in this study, performed as subexperts. They did not perform as the consulting experts did with the shortest solution path, but they were able to discern the significant bits of information. None of the live physician's performed as did NEOMYCIN. However, just like the residents in the Patel, Kaufman, and Evans (1989) work with live discourse, the medical residents in this study are more reliant upon getting additional and greater evidence from physical examination as they kept generating more and more hypotheses. It was because subject 2 was the most inexperienced resident that he or she did this quite ostensibly.
The first thing I think about is meningitis. And then I think about other things; that are more serious, but are rare.

...I would have take a more detailed history, basically and
<<How long has Debra had this kind of headache>>
(This would be a reasonable question to ask). You want to know how long she has had it. What kind of pain she is feeling and where is it exactly (is it going anywhere else in her head, or is it just staying there in the head going down to the neck). What else can give you a stiff neck? I mean. hypertension, bleed intracerebral bleed can lead to a stiff neck and a severe headache (that's usually in older people, but you still have to rule that out), she 39 years old. Blacks, usually they have higher incidence of hypertension, which would lead to intracerebral bleeds.

...Ah, there all different kinds of meningitis, you have encephalitis, also might ask if you'd ask if she's had previous infections, or any symptoms of the upper respiratory tract; she travel in an area endemic for encephalitis. Ah you would also worry about internal, Intracranial tumours
Ah with headache stiff-neck nausea and vomiting you'd also have to ask, trauma, ...fell down. got hurt, that could be it too. I mean this is just four simple little things here that could be in many many many other differentials.

Things, Drugs, you would have to worry about drugs. . You know. Those are the major things I would worry about. because... there are a whole bunch of other things, but you know there are very rare.

Except 27 shows how subject 2, in this study performs in a different way, only due to the different experimental circumstances than the resident in Kaufman (1987), but essentially does the same thing. Rather than focus on one symptom here (where an answer is not available from the patient), the subject talks about each disorder as a different path of inquiry. The internist subject 6, on the other hand (excerpt 1-4) structures his interview around the examination and history taking procedure to get to the bottom of the disorder. The internist, rather than talk about the many condition which can cause these symptoms allude to how to rule them out by history and examination. The subject behaved similarly to Kaufman's subject, generating hypotheses and focusing on symptoms rather than creating a working differential and proceeding from there. It should be noted, however, that the resident in Kaufman's study could ask all the questions desired, whereas subjects 2 could not have verbal
questions answered. The following table (5) indicates the initial diagnoses of the subjects in the experiments.
Table 5. Initial Diagnosis Upon Presentation of Presenting Complaint

<table>
<thead>
<tr>
<th>Subject</th>
<th>Initial Diagnosis Upon Presentation of Presenting Complaint</th>
</tr>
</thead>
</table>
| **Subject 1  **  
Resident            | Intracranial Problem: Infection (Meningitis or Encephalitis) or Space-Occupying Lesion |
| **Subject 2  **  
Resident            | Infection (Meningitis Encephalitis) Hypertension Intracerebral Bleed Intracranial Tumours Trauma Toxicity/Drugs Migraine |
| **Subject 3  **  
Resident            | Meningitis Stroke |
| **Subject 4  **  
Gerontologist        | Infectious Process: Meningeal Symptomatology Problem Pregnancy Hypertension Severe Headache |
| **Subject 5  **  
Gerontologist        | Tension Headache Migraine Infection Meningitis |
| **Subject 6  **  
Internist             | Infection: Meningitis Hypertension |
| **Subject 7  **  
Resident             | Migraine Cerebral Bleeding: Subarachnoid Haemorrhage Meningitis |
The man/machine interaction has some parallels to doctor patient dialogue. In doctor/patient interactions the patient is reporting his observations and thus presenting the physician with findings. Thus, the residents do not perform as parsimonious a physical examination and history and, in this study, they have more difficulty finding a differential they can rely on as much as the more experienced physicians.

In concordance with the Kaufman, Evans & Patel's (1989) study the subexperts engaged in considerably fewer exchanges with the patient, but were more accurate in their choice of question. From fewer doctor/patient exchanges, in that study, the experienced physicians asked not as many question, yet gathered as many positive findings as did the residents, and also encountered less negative findings. Thus, their inquiries were more focused and accurate, which implies that their hypotheses (or at least their internal models) were more accurate. The residents attempted to rule out or discriminate among various hypotheses (trying to build a model). Whereas, the more experienced doctors exhibited a greater capacity to acquire and manage patient information. In this experiment, the students are not free to interactively ask new questions, but they complain more about not having enough history and physical examination. Figure 8. might be interpreted as implying that the resident keeps building possible models making reference to the presenting complaint symptoms. On the other hand, the more advanced physicians, refined and elaborated hypotheses, refer back to their own differential diagnosis rather than stratagem from the presenting findings.
To sum up, the representation of experienced internist subject 6, which we
looked at the beginning of the chapter, indicate a unique point. This doctor, although
having a well ordered differential; generates many questions, and completes a
diagnostic interview in lieu of stimulus from the programme. Unlike the resident
(subject 2) the internist generates a complete clinical interview revolving around the
differential diagnosis to which he or she feels with certainty will solve the case. Again,
the nature of an internist's job is to do this, so that habit or tendency may be play major
factor in this physicians treatment of the interaction. Subjects 3 and 7, medical
residences make orderly differentials, but are not as certain of them as the experienced
physicians'.

ACADEMIC VERSUS DIAGNOSTIC CONSIDERATIONS:

Looking again at an experienced physician. Figure 12 a & b shows how subject 1,
an experienced physician, went about solving the case. The initial data on the patient
(headache, stiff neck on flexion, nausea and vomiting) lead the subject to immediately
consider meningitis, encephalitis, or a space-occupying lesion. The short duration of
the headache lead the physician to be less suspicious of a lesion and more of an acute
problem, like bleeding (subarachnoid haemorrhage). The absence of a fever put
infection further back in the differential. The exact meaning of the next question 6)
Has Debra experienced seizures recently? was not clear to this physician, who was
unsure about what part of the time line the recency fell into.

<table>
<thead>
<tr>
<th>excerpt 28</th>
<th>COMPUTER...Has Debra experienced seizures recently?</th>
</tr>
</thead>
<tbody>
<tr>
<td>SUBJECT 1</td>
<td>Again, I would want to know whether that was</td>
</tr>
<tr>
<td></td>
<td>recent; meaning since the onset of the headache,</td>
</tr>
<tr>
<td></td>
<td>or whether it's been in the more recent past -</td>
</tr>
<tr>
<td></td>
<td>before the headache started...ah...</td>
</tr>
</tbody>
</table>
Space-occupying lesion is less likely. It can be an infection, but the headache has not been for very long. → So the differential would go in the line of a vascular problem: such as a subarachnoid haemorrhage

2) Immediately upon complaint:

3) Headache duration: 6 hours:

Intracranial problem → either an infection: meningitis, encephalitis, or perhaps a space-occupying lesion.

4) Headache severity = very severe:

It may be a Subarachnoid Haemorrhage due to the severity of the headache. Infection is still on the differential.

5) Fever = none:

This does not rule out infection, but it moves it further back on the differential.

Figure 12a: Reasoning Considerations of Subject #1

— information made available from NEOMYCIN
6) Recent seizures; none:

Hæmorrhages can be associated with seizures, but seizures can be associated with things that bleed such as AVM’s and aneurysms and so on.

7) Visual problems; None:

One can get diplopia (double vision) with some types of space-occupying lesions, and also with some vascular problems you can have associated blindness (i.e. monocular blindness).

With the presenting story so far: Main Differential: Subarachnoid Hæmorrhage and visual problems are not associated with that.

8 Headache onset = Abrupt:

5) Funduscopic examination; Normal:

Main Differential: Hæmorrhages, as usually the symptoms begin very abruptly

The abnormal funduscopic exam gives one an indication of intracranial pressure. If one has a blurring of the (optic) disk margins, this is more suggestive of a space-occupying lesion that can either lead to a tumour... A large hæmorrhage can also give you that, but again, I expect the fundoscopic exam to be normal.

Figure 12b: Reasoning Considerations of Subject #1 (continued)

- information made available from NEOMYCIN
The question (6) **Has Debra experienced seizures recently?** is similar to (6i) Has Debra experienced seizures?: a question which implies pragmatically either:

(6ii) Has Debra experienced seizures recently, in connection with the present complaint? or

(6iii) Has Debra a recent history of seizures?,

with (6ii) being more appropriate, since her symptoms point to a condition which could cause seizures rather than seizures being a condition which could produce her symptoms. Thus it is, in part, the medical reality which drives the interpretation. Therefore, (6iv) Has Debra ever experienced seizures?; (6v) Has Debra a history of seizures? & (6vi) Has Debra a past history of seizures? are less relevant than even (6vii) Has Debra experienced seizures in the near past? It is interesting to note that in figure 9b question 6) (about whether Debra had any recent seizure and question 7) (as to whether she has any visual problems) are talked about in the subject's protocol not in the same way as other items which the doctor used to solved the case. Rather it was almost as if the physician had wondered aloud about why they were asked, in the first place, since they do not seem to fit in with the differential at this point: e.g., *"With the presenting story so far my main differential would be a subarachnoid hæmorrhage, and visual problems are not usually associated with that. "* That is, the subject made what is indicative of an academic association rather than one which would clear up or refine the differential. It is as though it were not a question that should be asked at that point. The consideration: **Hæmorrhages can be associated with seizures, but seizures can be associated with things that bleed** is circular and revolved around the subarachnoid hæmorrhage in the differential. Arteriovenous malformations (AVM's), which include aneurysms are the leading cause of subarachnoid hæmorrhages. When the information come about that the headache was abrupt, bleeding become more prominent. In the differential, the response to the visual problems question is more explicitly a comment about the perceived non-productiveness of asking whether the patient has visual
problems. The question put forth by NEOMYCIN was not ridiculous, but in light of the picture which this physician had in mind it did not fit in, and the subject did not expect to find positive results in that direction. The physician goes through a process of refining verbal data to get to the truth about the patient's real condition. What we have seen up until now points to two major assumption that professionals make, domain level assumptions with respect to their representations, and meta-level assumptions about interpersonal communication (or rather extrasemantic indices of interpretation). Each plays a part in the process.

Academic judgements, with reference to computer questions which the subjects deemed unrelated to their differential, like these, ruled out the concerns about the subjects being lead by the computer, rather than being active participants. Their objections to the opposite were ostensible: The physicians are trying to determine in short order "what this lady has". NEOMYCIN is "trying" to examine academic possibilities, notwithstanding whether such a speculation should be asked after her condition stabilizes.

The subjects, nevertheless, often indicated that they would have liked the machine to tell them things about medicine as a reference source, or reveal its reasoning on some criteria beyond the immediate medical consideration. They wanted to know this not only as a way of getting to know more about the condition of Debra, but also about related conditions and the considerations made by someone who sees problems like hers on a regular basis.
EXPERIMENT II

Most of the results for the second set of experiments refer back to the first experimental as a comparison.

EXPERIMENT iiA

These physicians, participants in experiment 1, asked their own questions, as they tried to solidify their differentials to make a final diagnosis. The physicians often restarted their interview asked question in a typical order of a differential diagnosis, of presenting illness, history, physical exam, and laboratory data as they explained their data requests. In this respect, they felt the need to review the case twice to be certain about their hypotheses that Debra either has a bleed or a meningitis, they acted like the expert consulting employed in this study to definitively diagnose the case.

In general, many of the conversational nuances that exist in experiment I disappear in experiment II, because the actions are performed by the subjects (subject 1,2,3,4,& 6) themselves, and the interpretation is no longer centered on projecting onto the programme. The subjects does not have to decipher what the computer doctor is doing. As noted, it was then, in this instance, that the findings themselves becomes more suspect. This is, the physicians desired here, even more than in experiment 1, to examine the patient physically themselves, and clear up what was missed by not asking questions the way they liked to ask question. They wanted to get to specific types of information from the patient verbal interview, physical examination, and laboratory report that was wanted (this was true, even for findings which were answered by the programme). The physicians also wished to ask other questions which the programme could not ask, but moreover, were interested in asking fuzzier, less straight-forward questions to refine the information about the nature of the findings. Findings which, they felt in real-life to be more subtle. Also, the physician wished to make up for
information lost by not actually being in physical contact with the patient, in experiment 1.

One physician discovered that the patient had a positive CT-scan for a subdural haematoma. It is noteworthy that a medical resident saw this same finding in the second experiment and did not find it so odd, possibly because it was assumed that a finding in the computer was absolute.

The physician knows that the data on physical examination and history do not fit with this laboratory finding. It does not mean however, that Debra does not have a subdural haematoma, but that would not account for her present symptoms.

Therefore the physician talks about the possibility of the haematoma being in a form that may have produced these symptoms, but still feels that there is something else which is causing the symptoms, not the haematoma.

We can see how certitude about the computer finding disappears in experiment II, when it is the subjects themselves who are asking the question to the machine. They wonder if the programme is answering what they intend and also they question its competence; e.g., "It doesn't say that Debra had a stiff-neck on examination. One
resident (subject 3) did not see the discontinuity of a subdural haematoma and the patients physical symptoms, but takes the results of the scan as a given.

This same subject also revealed a difference between how he or she responded with respect to verbal diagnostic requests, versus what was subsequently keyed into the computer in the form of requests. This was the only subject to have such a discrepancy. This physician stated an intention of looking for: aphasia, blood pressure, a CT-scan, headache-chronicity, headache-visual prodrome, headache location, headache episodic, headache exertion, head trauma sign, hemiplegia, hemiparesis, Kernig’s sign, sickle-cell, stiff-neck on rotation, and temporal tenderness. The subject eventually keyed in: headache chronicity, headache visual prodrome, Kernig’s sign, Brudzinski’s sign, head trauma sign, sickle cell, CT-scan, and CT-scan haematoma.

**EXPERIMENT IIB**

**DATA REQUESTS WITHOUT PRIOR EXPERIENCE**

Similar to figures 5a through 5f one can see from the schemata (figure 13) of subject 5 the nature of medical considerations which were developed by the physician in this interface. This subject is encountering the case for the first time and “running” the diagnostic inquiry, much as a doctor/patient investigation.

The figure schemata of medical reasoning employed in figure 13 has a similar convention to the schemata used to describe subject #4 in experiment I. There are some differences, however, and although the cognitive objectives do not differ between the two tasks, the fact that the subjects can inquiry about a finding, puts the doctor in the driver’s seat.

The square boxes contain information made available on the computer screen, and the rounded boxes enclose information originating from this experienced physician. The box labelled DECLARATIVE INTENT shows what the physician said he or she was intending to do. The ACTION: QUESTION box referred to the direct manipulation which the subject made on the computer, such as choosing the menu item
HEADACHE-ONSET to ask the question "What is the onset of Debra's headache?" The findings are the answers provided by NEOMYCIN to the question posed by the subjects menu choice. The box labelled SIGNIFICANCE contains reasoning and commentary with reference to the finding. One will notice, at this time, that the OBSERVATION boxes below the initial observation appear blank, and void of content. They are not. There is an arrow leading from the findings box on the previous line. It is these findings which are the new observations. Similarly there is a dotted line leading from the significance box on the previous line to the declarative intent box below. This is to denote that the physicians verbalization about the next procedure is made in light the perceived significance of the observed information. It is evident that the initial information which the doctors in both experiments see are the same: Debra is a 39 year old Negro lady presenting with a 6 hour headache, stiff neck, nausea, and vomiting. The difference here is that the computer displays its initial differential of meningitis. This holds true as well for the other part of experiment II.

It is interesting to note that this doctor explicitly talks about the procedure of differential diagnosis, stating that it was starting with vague general questions. The physician (subject 5- a gerontologist) asked about headache onset, and discovered that it was abrupt (this was the only physician who did not immediately verbalize a concern about bleeding). This was most likely due to the fact that this physician was only starting the history where it would be unwise to jump to a conclusion which could effect closure, whereas the other physicians in experiment I were considering the case for some time before they discovered the onset of the headache was abrupt. One can see, as the physician states, that a series of questions clarifying the nature of the headache are being put forth. However, due to a problem which arose, the programme failed to tell this physician about the level of severity of Debra’s headache.
OBSERVATION
39 year old Lady
Negro
Ruff Neck on flexion
Nausea
Vomiting
Headache
Adding
Differential meningitis

DECLARATIVE
INTENT
I want to start with vague
general questions to
localize her headache a
little bit, like when did it
start and that kind of
stuff

ACTION:QUESTION
HEADACHE-ONSET
FINDING
ABRUPT

SIGNIFICANCE
It doesn't tell me much
It eliminates chronic sort
of tension headaches or
something like migraine,
but not completely
because that started
suddenly, one day

OBSERVATION
DECLARATIVE
INTENT
It still looks pretty
general

SIGNIFICANCE
I guess they don't
know

OBSERVATION
DECLARATIVE
INTENT
Fever, or maybe I'll
do severity first so
I can finish my
headache questions

ACTION:QUESTION
HEADACHE-
LOCATION
FINDING
FRONTAL

SIGNIFICANCE
That's very interesting,
because if I am looking
for sort of infectious
causes of her headache,
it doesn't rule it out, but
it doesn't tell you what
it is either

OBSERVATION
DECLARATIVE
INTENT
I would like to know
whether she had some
kind of symptom like a
cough or a sore throat or
anything like that

ACTION:QUESTION
COUGH
FINDING
DEBRA DOES
NOT HAVE A
PRODUCTIVE
COUGH

SIGNIFICANCE

OBSERVATION
DECLARATIVE
INTENT
will ask about fever

ACTION:QUESTION
FEBRILE
FINDING
DOESN'T HAVE A
FEVER

OBSERVATION
DECLARATIVE
INTENT

ACTION:QUESTION
VISUAL-PROBLEM
FINDING
NO VISUAL PROBLEM

SIGNIFICANCE

Figure 13 Schemata of Clinical Reasoning of Gerontologist
Subject #5
The assumption on the part of the physician was that the programme did not know. This is, of course, a very reasonable assumption, however when GUIDON does not know something it clearly says unknown. It would be pragmatically reasonable to take silence as an index of ignorance, the way we do in people.

A crucial finding like fever makes quite a difference in a differential, the fact that Debra does not have a fever has different effects on the reasoning of the physicians. This doctor, in experiment IIb, makes the most eloquent statement about the absence of a fever in a patient one could safely assume has a high probability of having an infection.

Subject #4 (see schema figure 5c) used fever in a lot of “rules” e.g. The physician still could be considering meningitis in spite of no fever and no seizures. She has a migraine a very bad headache which won’t go away or viral meningitis. Subject 4 considered it with infectious process, hypertensive crisis less likely

The network on subject 4 (figure 5c) shows that the fever is linked with all of the hypotheses in the differential indicating that this physician at this point has not made a clear-cut diagnosis. Looking again at the network of subject 1 (figure 7), evidence of fever here pushed infection further back in the differential. Seen more clearly in figure 12, the reasoning consideration were stated, “It (the absence of fever) does not rule out infection, but it moves it further back on the differential”. The medical resident - subject 2 (figure 8) kept meningitis high in the differential.

Continuing with the results on the subject in the second study, one can postulate that the question about cough could be to indicate an upper respiratory infection.
In the end this subject, as the others, suspected meningitis and subarachnoid hæmorrhage and considered migraine, but this physician was having difficulty figuring out what was wrong with the patient, and did much back-tracking.

**EXPERIMENT IIC**

This experiment allowed the subjects to take a self-directed exploration of the system. An interesting item of note is that one doctor looked at the graph-conceptual representation of the diagnostic process of NEOMYCIN called the Patient-Specific Model of Debra. The physician wished to add to that model a new hypothesis to the differential and a new finding referring to the result of a CT-scan which he or she had asked for. The semantics of the rule produced unfortunately did not reflect what the doctor had intended.
So, with respect to natural language and computer translations for physicians who would wish to modify the knowledge base for their purposes, one can see it is not easy to translate semantics even with a very friendly system such as NEOMYCIN has.

Wilkins (1989) performed a study on an aspect of NEOMYCIN in order to enhance and refine its knowledge-base for cases which it did not have the rules to solve. The methodology of using the part of the programme (the ODYSSEUS MODELER which can modify knowledge bases for HERACLES) as a way to enhance the problem-solving ability of the programme, as well as enhance its ability to explain and thus instruct. Swartout & Smoliar (1989) suggest that the same knowledge base used to provide clear explanation to users in a domain should be used to solve the problem. This constraint makes the programme representationally explainable. Swartout & Smoliar (1989) indicate that a knowledge base's ability to explain itself to users in a domain should be the criteria used in its construction. So the the same representation that drives the explanation drives the inference engine.

They propose a method of programming that would lend ease and utility to future user modification. Causal relations should be decomposed and not simply refer to on computer 'concept' because these are inassociative 'meta-concepts', which cannot be decomposed to relate to other 'concepts', e.g., [SMOKING-THAT-LEADS-TO-LUNG-
CANCER). Swartout & Smollar specify that the programme lines should be written as relations, e.g., CAUSES(SMOKING, LUNG-CANCER) if the expert virtually always talks about the relation as a simple link without qualification, and as a reflected "concepts with relations, e.g., [CAUSE-47 (specializes CAUSE)

CAUSER: Smoking

CAUSEE: LUNG-CANCER] when experts typically qualify and elaborate upon the relationships. This might be one way for systems to fulfill subject expectations more often.

In summation, the series of experiments indicated that subject expectations, and subtle phrasing have an effect on human understanding of the interface. Moreover, the perspective of the user affects how the interface is interpreted. However, not all interpretations are equally valid, and can be contradicted by prior or subsequent information. The structuring of information, keeping these considerations in mind, no doubt, could potentially reduce unsubstantiated inferencing. Some of these assumptions affect the course of problem-solving, thus, they should not be overlooked in teaching, or professional settings.
CHAPTER VII

CONCLUSIONS

In review, this thesis has consisted of a series of experiments in which physicians have interacted with an intelligent tutoring system. The subjects, who were observing a medical case unfold in a computer environment, were urged to think aloud as they themselves were working through the case. The experiments were developed in order to monitor inferential assumptions which physicians made in respect to attributing intent and purposeful behaviour to the activities of a programme as well as to the linguistic interaction of a man/machine interface.

The subjects were also told to make explicit certain other considerations beyond simply thinking-aloud. They were asked to make explanations, predictions, and to make a critique any element of the circumstance in which they were. Halliday and Hasan's (1976) indices and analysis of questions has proven to be a useful way to categorize responses, but it was necessary to modify them for the purpose of this study. This is because responses, to the experimental on-line probes for the subjects to think aloud, and reactions to the discourse on the computer display can to be thought of as responses to questions. The outcome of this study indicated that subjects were able to perform the experimental tasks meaningfully, while working through the case. This also indicated that experiments, such as this, are viable ways to monitor general, domain-specific, and conversational inference with respect to problem-solving.

The results in the previous chapter suggest answers to the questions and hypotheses posed in Chapter 3, and we will now evaluate the conclusions drawn from the experiments.

1. Are the physician's subjective interpretations of the linguistic exchange so extensive as to constitute anthropomorphisation?
Yes. Moreover, the doctors do distinguish between characters and their roles in the interface, rather than treat the interface like a 'canned' demonstration programme. The process of anthropomorphisation was found to have a significant impact on the subjects’ reasoning. This was demonstrated in the section entitled 'anthropomorphisation' in the results. (page 98) This finding is so noteworthy it leads us to the findings related to the second question.

2. Are there similarities between the subjects' interactions with the programme and their interactions with people?

Yes, especially with respect to the anthropomorphised roles which the computer personages play in real life (see page 105).

It can be seen how the subjects could have had difficulty determining the source of the patient information. In the beginning of the case, the findings provided in the computer's question and answer exchange, it would have been safe to assume that the information was coming from 'Debra' (for this is where history data usually come from, assuming the patient is conscious). Since this mirrors real life, most subjects did not question it. In the middle of the case, when medical signs were being requested, and towards the end, when the laboratory results came in, it became clearer that other medical professionals would be the sources of this information. The question of who takes over, and at what point, during the transition between these stages may be the source of some of the confusion. Above and beyond this, some subjects may have been dealing with the programme as though they were reading of a set of cue-cards, but were wary about how the information initially got into the data base. Specifically they may have been concerned as to whether the "real" patient was giving reliable information, and the "real" doctors were trustworthy in their findings. Moreover, one may wonder if this machine was telling them the whole story. It is a strange phenomenon that when people cannot understand a story they will sometimes blame the story-teller.
The quality of the on-line think aloud verbal protocols indicated that presenting dialogue text, which is similar to an exchange between a doctor and a patient (in a question and answer form), induces physicians to discuss matters which pertain to human-computer interface.

Most of the subjects were more mistrustful of the information they felt was emanating from the "patient" than from the "doctor". If a 'physician' was perceived as being competent, then in an instance when the 'physician' appears to have overlooked an item or a procedure in the diagnostic interview, that item is seen almost as unimportant. One the other hand, if the "physician's" report and strategy is not deemed to be trustworthy, any apparent oversight is seen as a mark of incompetence. In contrast, when subjects empathized with a computer 'patient', the 'patient' reports were taken at face value. Otherwise, the accuracy and reliability of the information conveyed was treated with mistrust. The subjects even addressed the issue that there was a 'doctor' who saw the patient before the current 'doctor'. Possibly, they interpreted the interface as an on-line transcript of an actual doctor/patient exchange.

As to whether or not the patient data can be trusted, the subject-programme interaction in which this consideration arose, took place, mainly, at the level of the subject's encounters with the COMPUTER FINDINGS. The interaction, pertaining to the validity of the doctor to 'doctor' consultation, took place at the level of the COMPUTER EXPLANATIONS.

There were some failed subject expectations in the interface, such as not being able to get more information about patient findings in the COMPUTER EXPLANATIONS, and in not getting reference source information about what would constitute normal levels in blood and cerebrospinal fluid, as no on. The subjects sometimes found NEOMYCIN's considerations and actions to be too academic (e.g., the instance where NEOMYCIN asks whether or not Debra had polycystic kidney disease). It was felt that these considerations were not helpful in making decisions about relieving the present
emergency The fact is that NEOMYCIN was not in an emergency mode, while the doctors were. In addition, the subjects sometimes found the programme's explanations did not always answer the questions which they had assumed they had asked. For example, GUIDON usually explained its course of action, in relation to the most proximate rule-driven relationship which triggered the move, rather than stating why it was doing a particular action at that time, instead of other strategies. That is, it did not provide the contextual rationale behind its reasoning. In this way it did not respond with all the pragmatic subtlety of a human being. (i.e. in its recognition of intent).

There are other parallels between man/machine and person to person interaction. Doctor/patient discourse is an interpersonal speech event arranged around the exchange of information. The clinical interview is the medium through which most physicians get information from patients. Most diagnoses are made during the history-taking (Kassirer & Gorry, 1978). Interviews are structured, predictable and organized around specific topics. In an interview, patients can demonstrate competence or incompetence in answering questions (Fisher, 1983), as they were perceived to in this study. However, in this study the patient could not ask a question, the 'doctor' could not clarify a finding, and the user's reflections, questions, and queries about the instructions and task — subjects which could be cleared up easily in human/human interactions — were not available to the machine. The responses to the questions displayed by the programme, as we have seen, lends to giving careful consideration to the semantics of questions and the presuppositions which are inherent in WH-questions (Belnap & Steel, 1976). The effect that NEOMYCIN's questions had on the subjects was not only semantically-base, but also paralinguistic. That is, their interpretation of the strings depended upon the cognitive perspective in which the subjects were. The nature of their problem-space of the case (especially with
regards the role of the computer personage) when they encountered the lexical string, also played a part in their understanding of the exchange.

3. Does the experience of the user in the domain of expertise have an ostensible presence in this study, indicating different kinds of reasoning at different level of expertise?

Yes and no. Only the more experienced subjects were, in this experiment, considered to be low domain experts, with the exception of the residents. However, the case was not above the level of complexity which a resident could handle. Virtually all the of subjects, therefore, functioned as subexperts and thus, performed at about the same level. However, one subject, just beginning residency, did indicate novice-like clinical interview behaviour (page 141). The subjects who possessed higher levels of expertise had the ability to form hypotheses with less of data. These physicians were less distracted by the novelty of diagnosing on a computer, and showed different types of clinical reasoning.

The purpose of this study, however, was not to examine expert/novice differences, but since differences did occur, they were duly noted. The novice behaviour was indicated by the subject's inability to elaborate or to develop the hypotheses initially made, towards a more specific diagnosis. Instead the subject continued to generate additional hypotheses in response to new data (see ). The real experts, employed in this study as consultants, did not perform within the experiment, and thus there was no way to predict their likely performance.

4. Did the task of explaining, predicting, and critiquing actually play a role in the decision which the subjects made, or were they unnatural tasks which confused the subjects?
The results of the study (page 90) indicated that subject's reasoning was closely tied to subject explanations, and that critique and anthropomorphization also were paired. The relationship between reasoning and explanation is that the latter is the vocalization of the former. The association of critique and anthropomorphization is that subjects spent most of their time making criticisms directed toward the reasoning and the nature of the case, rather than towards the environment, so that the criticisms were directed at the personages, and at their intent and reasoning.

In the interaction between the subjects and the programme the cognitive tasks of explanation, and critiquing, substantially influenced their diagnostic reasoning.

Explanation did not appear to be an unnatural task which interfered with medical reasoning. The subjects were not accustomed to explaining each move they made, but they quickly adapted to it. This is not surprising, as physicians are routinely required to engage in explanatory behaviour. One can reasonably assume that there were many considerations which they did not verbalize.

Explanation forces the subjects to think about their strategies and of those of the programme, and critiquing makes their evaluation of information and procedures explicit.

Critiquing did seem to be more of a natural task and was a medium which made explicit other implicit considerations explicit which may not have been captured otherwise with explanations. Critiquing produced rich protocols. During the experimental consultation, the subjects' critiques focused on different aspects of the interface: They made critiques about the computer's medical questions and its replies with respect to its syntactic form. The use: (the intent and application of those propositions) were also critiqued by the subjects. Likewise they critiqued the programme's content, which is the feasibility of its medical reasoning. GUIDON's 'explanation' facility transformed the act of critiquing into an interactive pseudolinguistic exchange between the subject and the programme. The more the
programme explained itself, the more the subjects personified it. Explanation and critiquing have proven to be viable and interesting windows into reasoning. They allowed one to examine not only what was being considered but what is not being considered in terms of making rational presuppositions more explicit.

Prediction was the least natural task. What the subjects ended up doing was talking about what they themselves would do next. Prediction was not a very strong category since subjects did not do much active prediction, but responded to this probe by keeping their current and future operations in mind. Subjects, generally, did not predict potential computer findings or questions per se, each time a sentence string appeared on the screen. What they predicted ended up being implied in other probes. Their prediction were embedded in the presumptions made, when they made explanations, personifications and critiquings. Most “predictions” were subsumed within these other manifestations of subject reasoning. The subjects made inferences which may appear to non-medical people as pure leaps of faith, such as assuming intracerebral bleeding from an abrupt headache.

- The subjects did function as active participants running their own diagnoses parallel to the programme.

- Doctors sometimes found NEOMYCIN too academic in its explanations and in some of the questions it asked. However, the programme was being used in atypical way, and the subjects were not using the programme to learn a slice of clinical neurology.

- The physicians did not respond to the case as though they were merely given an exercise. It is not known, for certain, whether subjects felt that they were being tested except in this instance when subject 6 said: I wonder who programmed this? Eh? as this physician looked at the experimenter smiling.

166
This same doctor had previously commented:

**excerpt 32**  
**SUBJECT 6**  
She has a no. Ha ha we have a tough case here! What it tells me... She's had no fever. She's had no seizures. She has a normal funduscopic examination. So, I presume that she doesn't have hypertension, and she hasn't got any brain tumours. She's only had the headache for six hours.

Otherwise, the subjects did not seem to perceive NEOMYCIN as a test of their competency. This doctor's assumption, as the others, is based upon a belief that the case is valid, an assumption that would lead, not to its difficulty, but rather to its impossibility. Subjects were more critical of the "doctor" than the case.

**LIMITATIONS OF THE STUDY**

One major limitation of this study is that it is mainly descriptive. It is required that one to find more explicit and accurate method of analyses. One that could arise would be something between a story grammar or prose analysis (Meyer, 1985) a possible set of loosely binding rules for erotetic logic, and usage conventions. I would tend away from the logical or propositional rules, however, due to the many degrees of freedom in social discourse, some of which are not very regular, and because the conventions in studies like this are going to be domain-specific.

Even though the investigation was very specific, the analysis was exploratory, and the categories of reasoning need to be examined more closely, in order to see what role they play in instruction (both in teaching and learning) and medical problem-solving. The task did not interfere intrinsically with the problem-solving, it simply changed how information was used in the subjects' problem-solving, and how the subjects talked about it.

The methodology combined the assignment of a procedure (which called for medical reasoning, but which does not interfere with problem-solving) and a structured interview, to which the subjects answered freely. The subjects were not verbally cued each time they saw a new stimulus. The doctors became accustomed to talking aloud about information in this way when they were in the pre-experimental instructional
trial. Although this method did produce rich verbal protocols, it is still difficult to separate out the proportion of information peculiar to the task, and to their general problem-solving.

Retrospective reporting was not used in this study, so that subjects looking at a videotape of their performance could comment as to what they were thinking at that time. This technique might have revealed more information about the logistics of the situation, such as their saying, "This is where I thought the task was too obvious, or too easy, so I choose something rare." (R. Tamblyn — personal communication). Although it was tempting to do both on-line and summary protocol analyses (Elstein, Shulman & Sprafka, in press), it would have been difficult, however, to persuade subjects to listen to two hours of their own audiotapes and to make a running commentary on them. It would have meant listening to them for several hours.

NEOMYCIN is only one example of several ITS environment possessing a microworld, again the experiments used this programme in an atypical way, which its developers did not have in mind during its construction. Nevertheless the subjects did respond in a consistent way to the conversational aspects of the interface.

IMPLICATIONS

Action theory (van Dijk & Kintsch, 1983), shows that at least part of comprehension subsumes goal-oriented, intentional, conscious and controlled behaviour. Since the subject expected "intelligent" (in the sense that it would perform like a human being) purposive behaviour from a machine which could "diagnose". Then anything which it seemed to do towards this end was seen as being meaningful. The subjects assumed that the programme had been consciously programmed for each step it took, and thus it had "intended" all of the specific steps it took as a means to consequences which were foreseeable by the programme. That is, whatever intermediary step which NEOMYCIN took, that subject-interpreted step was seen in the context of an instance to solve the case of Debra. Thus there were no instance when if
the programme did something unexpected and someone remarked, "Hey guy, your machine's busted!" Subjects, rather, commented that it was going about finding the disorder in the wrong manner, not that it was buggy and that the move was meaningless. Human intelligence perceives the purpose of things, and reasoning draws consequences as to the use of the means to obtain the ends which the subjects expressed, via their language, along with feelings and ideas. The subjects, were dealing, in part, with a 'black box' trying to determine why GUIDON was 'behaving' a particular way. For conversational interactions, as present in the exchange subjects had with the programme, however, there is more than one actor planning intermediary steps (sentence by sentence) to get to the whole "meaning" of the conversation (text). Meaning, however, is not context, but one uses context in order to understand things, especially things about which we have insufficient information.

The experimental situation was new for the subjects in which they assumed cooperative behaviour on the part of NEOMYCIN as they tried to figure out the rules of the game. One could apply a model of usage conventions, which could show how subtle phrasing differences could lead the interpreter to come to different conclusions about the meaning of an utterance (Hintikka, 1976). Deriving such convention for a particular domain may reveal sociological, and sociolinguistic elements in the nature of the social intercourse in that field. Together with identifying attitudinal and intellectual biases, it could also point out heuristics used by professionals to manage information in their work.

If the subjects are unconsciously making considerations about the likely and possible interactive effects of the assumed goal directed behaviour of NEOMYCIN, it is this that leads them to frame higher order intentions about the desires and beliefs of the programme - thus assuming rationality (Dennett, 1981). They do not have to assume that this is its thinking per se, but that its programme is goal-oriented and logical. In clinical training, as in all high level instruction, nothing is done just for its
own sake. Students learn how to do an interview, and to question findings so as to be more clear about what they observe. The same follows for learning to solve medical cases in an ITS environment. When working in a simulation of a medical problems nothing hazardous can occur as in real life (where people can die because of mistakes). But in a microworld, context need not be stripped, so that the context may reveal the consequence of an act. For pedagogical purposes, one may implement cost and time constraints to the patient model. The patient case can be made to evolve the disorder chronologically in real time, so that the order and timing of the user’s actions play a part in the case, as in ILE’s. The whole issue revolves around how medical people perceive these situations, whether they be simulated or real.

The subjects in this set of experiments had constraints; they did not have unlimited time (due to other real appointments) and they do not have access to everything about the patient, so as to be able to verify all items of information. So they are compelled to spot check (for that is what the system gives them) and to extrapolate from that limited information. This is the accurate and efficient discriminating of suppositions from seemingly limited data which is a mark of expertise. These inferences may have lead them to think (consciously or not) of possible worlds, or situations in which they needed to make assumptions. Thus, each possible patient situation concocted, for the constellation of symptoms, may entail a different course of action.

Most instances of misunderstanding did not arise as a result of semantic ambiguity in NEOMYCIN’s phrasing, but rather due to conversational atmosphere generated by the format of this experiment.

The study demonstrated that the consideration of pragmatic elements can provide insight into the nature of human/computer interaction. This merits further investigation in particular it would be useful to focus on other aspects of pragmatics with reference to semantics and comprehension in reasoning tasks. Also there is no
reason to believe that the kind of results observed in this study are specific to the
domain of medicine, or even machine interaction. The present information processing
Zeitgeist leans towards having the machine do most of the task, but as we can see it is
important that we focus on the nature of people in their interaction with machines.

Making explicit what is implicit in a complex task is one way to teach strategies
in complex learning. Medical schools and teaching hospitals have been moving to
change the traditional didactic method of medical education. The didactic curriculum
consists of teaching basic science courses and then pathology, essentially introducing
students to clinical experience late in their initial medical education. The change has
been in the direction of problem-based instruction (Schmidt & DeVolder, 1984),
wherein the students are introduced to real-life, and to simulated cases early in their
medical training. The rationale for this sort of curriculum is to build up the students' clinical expertise sooner and more robustly in their early study. The goal is to integrate
the medical information medical students are learning more tightly within the context
of the domain. Computer technology is one medium proposed to allow students to have
earlier opportunities for interactive learning, without the risk of patients being used as
guinea pigs. What this study has demonstrated that it is important to investigate how
persons understand information beyond direct semantic interpretation. In other
words, how the student, and the professional interpret their experiences, what subtle
beliefs are being held. It was noted that when the subjects asked the computer question
themselves, although they still personify the interface, they make less inferences about
why things happened the way they did. Differences in diagnoses or in the procedures of
clinical reasoning might be just as much related to assumptions that physicians make
with respect to the facts, as to the facts themselves. Miscommunication can occur upon
reading the wrong thing into a person's expression—(as in projection). It can also
happen by missing the message sent. As well in a human/computer interface
miscommunication on the part is not always due to semantic understanding, but can
also be the result preconceived ideas on the behalf of the recipient of information. One of the goals of instruction is to make these assumptions educated.

Again, in terms of pedagogy and the nature of these studies, (Winne, 1989) stipulates that a workable theory of learning for the development of ITS’s should move beyond the computer working as the programmer’s surrogate. The computer application cannot be programmed only to display information. It must be an environment which stimulates students to learn by forcing them to think by cueing them. Much of what affected the doctors in this study was the pedagogic explanations and questions which the developers of NEOMYCIN put forth, with the intent of stimulating students to think about their medical problem-solving strategies. As the subjects were subexperts, when they did run into things of which they did not have much understanding, they were forced to think about what they were doing, and what the programme was doing, and in some instances they did indeed learn new things. No subject walked away from the experiment without commenting about how interesting it was. This is noteworthy since two hours is plenty of time to take away from a practising physician.

Winne’s (1989) model of instruction has three parts. 1) through the programme the programmer knows something which the subject does not, and 2) intends to explain what he or she knows that the subject does not. 3) the subject must intend to comprehend the information from that explanation. In the context of this experiment, NEOMYCIN’s instructional academic cueing and the experimental probes served to make the subjects think about the information (both procedural, and declarative) in a more active manner. One way for this cueing to work effectively is that it must adapt to the user. This is possible in NEOMYCIN which can answer different types of explanations in its other modes.

Of importance also is the role of affect (feeling or mood) upon cognition. Much of the conversational effects were ones of attitude. Winne (1989) remarked that when one
noted variable, and unstable performance that cannot be explained by cognitive strategies then affective variables might be playing a role in the performance.

CLOSING REMARKS

One goal for the use of computers in instruction is to have a mixture of applied and basic research that is relevant to education practice. However, if one treats only a part and builds a theory to describe the whole, one is apt to be mistaken. Thus, practitioners should be guarded about generalizing from a specific piece of research which present panaceaic solutions (Kerlinger, 1977) In addition, computerized diagnostic reasoning courses should be practical and evaluation relevant to clinical practice performance (Dykes, 1981).

One may ask what interest this piece of research has for the physicians in the study? Will the outcome of this research have any importance in their lives? Subjects indeed, were highly motivated to come and work with the machine, but they did not feel that such a machine would do much for them in their practices, except to store information and act as a reference source. In this respect, they felt it could tell them things they did not know and would be useful for rare and difficult cases which they did not see in their everyday practice, but they certainly did not want to have the machine as a decision maker for them; unless the information such technology provided in the future is sufficiently reliable (Seigel & Parrino, 1988). In terms of the interface, besides making the system user-friendly so that users do not have to possess sophisticated knowledge to use the technology, users are less likely to believe the conclusions reached by an expert system unless it can justify or explain its conclusions in an understandable and convincing way (Zeidner, 1987). Typically physicians feel that if you cannot do something for them and their patients, you should not waste their time, as the average physicians feel overburdened by their practice. The challenge for this type of research which pertains to psychology, education, and interface design work is that it has also to be relevant towards patient care, and to giving the practitioner more
time, and to enhancing their efficiency (Lewis, 1983). The fact that subtle things about our belief systems affect our reasoning is an issue for education and practice.

The final issue, which is the theme of this work, is that in creating a user interface in this technology there have been conscious attempts in designing the intentional face — the end with which the public interacts. For example, some systems are intended to represent the user in the form of a Socratic teacher and others, in the form of a qualitative simulation. The fact is that AI-based instructional systems vary from simulative reification, to directive tutoring, from traditional problem-solving environments to simulations for teacher training (Wenger, 1987).

Pertaining to human/computer interface, authors raise particular concerns and have elucidated certain empirical findings. Winograd and Flores, (1986) propose having a computer system for many users which performs as though it understood the illocutionary force of the speech acts of requesting, promising, offering, accepting, reporting and acknowledging. Shneiderman, (1987) warns about adding too many features, options, and commands which overwhelm a user, and points out that natural language interfaces, by themselves do not provide users with information about actions and objects in the task domain, and operate as if to assume that the user is familiar with the task domain. Suchman, (1987) indicated how responsiveness to a system is limited to those occasions where the user’s actions affect some change in the machine’s state. Greenblatt & Waxman (1978) found that novice users were far more successful in retrieving data from a spatially-formatted presentation than a verbal one.

All of these and other design issues are what produce the resultant friendliness which makes certain computer systems easy to use even for a novice. Each system is deliberately constructed in a manner so as to personify the interface in particular ways. The personification may or may not correspond to the stated intent of the designer. It would be of value therefore to operationalize anthropomorphisation as a psychological
construct. This allows us to observe and evaluate, from instance to instance, what are the important considerations are, and why they are so.
REFERENCES


Kerlinger F. N. The influence of research on educational practice (1977, November) *Educational Researcher*.


182


Saussure, F de (1986) *Course in general linguistics*. C. Bally, A. et al. (Eds.) Lasalle, Ill.: Open Court


183


184
APPENDIX A

Patient Case 1000

2-mar-87: 08: 14
(Consultation of 21-Oct-85 10: 51: 28)

Please enter information about the patient.
This patient is a 39 year old woman who has a 6 hour history of stiff neck associated with nausea and vomiting. (Not all subjects saw this string since it went by too fast.)

2-mar-87: 08: 14
(Consultation of 21-Oct-85 10: 51: 28)

Please enter information about the patient.

<table>
<thead>
<tr>
<th>Name</th>
<th>Age</th>
<th>Sex</th>
<th>Race</th>
</tr>
</thead>
<tbody>
<tr>
<td>Debra</td>
<td>39 YEARS</td>
<td>FEMALE</td>
<td>BLACK</td>
</tr>
</tbody>
</table>

2) Please describe the chief complaints:
**HEADACHE
**STIFF-NECK-ON-FLEXION
**NAUSEA
**VOMITING

3) How long has Debra had this kind of headache?
**6 HOURS

4) How severe is Debra's headache (on a scale of 0 to 4 with 0 for very mild and 4 for very severe)?
**4

5) Does Debra have a fever?
**NO

6) Has Debra experienced seizures recently?
**NO

7) Does Debra have visual problems?
**NO

8) How would you characterize the onset of Debra's headache (Abrupt, rather SUDDEN, but not abrupt; or GRADUAL)?
**ABRUPT

9) Does Debra have an abnormal funduscopic exam?
**NO

10) Do Debra's headaches have precipitating or aggravating factors?
**NO

11) Does Debra have an abnormal mental status?
**NO

12) Does Debra have a history of polycystic kidney disease?
**NO

13) Does Debra have a family history of polycystic kidney disease?
**NO

14) Has Debra taken medications recently?
**YES

15) You mentioned that Debra has taken medication recently. Please elaborate:
**ANTIMICROBIALRX
**ANALGESICS

16) Has Debra been hospitalized recently?
**NO

17) Do you know the results of a recent CBC of Debra?
Continuation of TABLE

** YES

18) What was Debra's peripheral white count (in thousands)?
** 13

19) Percent of WBC's which were PMN's:
** 50%

20) Percent of the peripheral WBC's which are immature forms:
** 12%

21) Has a lumbar puncture been performed on Debra?
** YES

22) What is the total WBC count of the CSF?
** 330/MM3

23) What percent of the WBC's in the CSF were poly's?
** 55%

24) What is the protein value (mg/100cc) in the CSF?
** 635MG/DL

25) What is the glucose value (milligram percent) in the CSF?
** 53MG/DL

26) What is the simultaneous blood glucose?
** 110MG/DL

27) Is the CSF bloody?
** YES

28) Has Debra been Exposed to any contagious disease recently (e.g. meningococcal disease, mumps)?
** UNKNOWN

29) Does Debra have any cutaneous lesions or rash on physical examination?
** NO

Final diagnosis:

<table>
<thead>
<tr>
<th>HYPOTHESES</th>
<th>BELIEF</th>
</tr>
</thead>
<tbody>
<tr>
<td>INFECTIOUS PROCESS</td>
<td>+++</td>
</tr>
<tr>
<td>MENINGITIS</td>
<td>+++</td>
</tr>
<tr>
<td>ACUTE-MENINGITIS</td>
<td>+++</td>
</tr>
<tr>
<td>ACUTE-BACTERIAL-MENINGITIS</td>
<td>+++</td>
</tr>
<tr>
<td>VIRAL MENINGITIS</td>
<td>++</td>
</tr>
<tr>
<td>CHRONIC-MENINGITIS</td>
<td>+</td>
</tr>
<tr>
<td>SUBARACHNOID-HAEMORRHAGE</td>
<td>+++</td>
</tr>
<tr>
<td>BRAIN-ANEURYSM</td>
<td>+</td>
</tr>
</tbody>
</table>

or

Final diagnosis:

<table>
<thead>
<tr>
<th>HYPOTHESIS</th>
<th>OF</th>
<th>CUMCF</th>
</tr>
</thead>
<tbody>
<tr>
<td>INFECTIOUS PROCESS</td>
<td>500</td>
<td>933</td>
</tr>
<tr>
<td>MENINGITIS</td>
<td>500</td>
<td>987</td>
</tr>
<tr>
<td>ACUTE-MENINGITIS</td>
<td>0</td>
<td>974</td>
</tr>
<tr>
<td>ACUTE-BACTERIAL-MENINGITIS</td>
<td>974</td>
<td></td>
</tr>
<tr>
<td>VIRAL MENINGITIS</td>
<td>524</td>
<td></td>
</tr>
<tr>
<td>CHRONIC-MENINGITIS</td>
<td>400</td>
<td></td>
</tr>
<tr>
<td>SUBARACHNOID-HAEMORRHAGE</td>
<td></td>
<td>989</td>
</tr>
<tr>
<td>BRAIN-ANEURYSM</td>
<td>444</td>
<td></td>
</tr>
</tbody>
</table>

187
APPENDIX B

DETAILED INSTRUCTIONS AND EXPERIMENTER PROTOCOL

If the physician, at this point, has agreed to be a subject he is asked the following:

INSTRUCTION 1: "You will then see the the CONSULTATION window; in it you will see something like the following appear on the screen."

INSTRUCTION

'This is a CONSULTATION of a particular PATIENT CASE.

COMPUTER INFORMATION ON PATIENT CASE 1001 - DIANE.

COMPUTER 1001 DATE TIME

(Consultation of 4-Jul-83 10: 12: 00)

----------- Patient 1001 ------------

COMPUTER

"PLEASE ENTER INFORMATION about PATIENT CASE 1001 Diane"

"EXPERIMENTER-EXAMPLE.

"PLEASE ENTER INFORMATION about PATIENT CASE 1001 Diane".

INSTRUCTION

It contains patient information such as the name, age, sex and race;

COMPUTER

Diane is a 42 year old woman with a headache for the past eight days.

EXPERIMENTER-EXAMPLE

'Diane is a 42 year old woman with a headache for the past eight days'.

EXPERIMENTER-EXAMPLE

Question # 1) The patients chief complaints

COMPUTER NAME AGE SEX RACE

COMPUTER 1) ** Diane 42 Female Latino

EXPERIMENTER-EXAMPLE

'Diane is a 42 year female Latino'.

INSTRUCTION

1 (Nota Bene:

COMPUTER will refer to what is displayed on the Computer screen.

EXPERIMENTER-EXAMPLE will refer to what the experimenter is saying; often—times he or she is reading aloud the computer display and acting as though he or she was an experimental subject. This is to show the subject what to do.

INSTRUCTION refers to the explicit instructions that the subject will be given during the demonstration).
Question # 2): The first question about the patient that the programme is asking.

COMPUTER

2) Please describe the chief complaints:

COMPUTER       ** HEADACHE
COMPUTER       ** STIFF-NECK-ON-FLEXION
COMPUTER       ** NAUSEA
COMPUTER       **

Note: The experimenter continues...

EXPERIMENTER-EXAMPLE  "We know that she is complaining of a headache, stiff-neck-on-flexion, and has been experiencing nausea".

INSTRUCTION

A) "Please continually read aloud the patient case information; say aloud your thoughts as each question is presented. Also indicate what the answer to the current question tells you about the patient case. Each time a new piece of information comes on the screen, please indicate aloud your thinking at that point. What does the present line of inquiry contribute to the case?"

EXPERIMENTER-EXAMPLE

Question # 3)

COMPUTER

3) How long has Diane had this kind of headache?

EXPERIMENTER-EXAMPLE

"How long has Diane had this kind of headache?"

INSTRUCTION

B) "After this, and at all times, please predict what the next question or group of questions might be; e. g. 'Does Diane have persistent headaches of this kind'"

INSTRUCTION

C) "Also please attempt at each pause to predict what you feel would be likely responses to those questions; in other words, predict what some possible answers to these questions might be'. DO NOT CONTINUE UNTIL YOU HAVE FINISHED INDICATING YOUR THOUGHTS ABOUT THE NEXT QUESTION AND ANSWER.

INSTRUCTION

D) "Please, as I have done read aloud the patient case information; say aloud your thoughts as each question is presented. Indicate what the answer to the previous question tells you about the patient case. Each time a new piece of information comes on the screen, pause to indicate aloud your thinking at that point. What does the present line of inquiry contribute to the case?"

INSTRUCTION

E) "Now the programme will pause. A little menu to the top left of the CONSULTATION window will appear. Using this menu you will have the choice to either have the programme 'CONTINUE' (in order to attain the answer to the programme's current question, and to simultaneously allow it to generate the next question) or you may choose EXPLAIN — to have the programme explain the reason for its present question. To operate this menu bring the arrow to the menu choice with the computer mouse that I will show you and depress and release the middle button on the mouse. The choice will turn
black as you depress. The subsequent release of the middle button completes the selection.

EXPERIMENTER-EXAMPLE

"I will choose CONTINUE here", and I will indicate where on the screen I choose it.

NOTE

The experimenter will select continue and you will see the answer to question #3.

COMPUTER

**8 days

EXPERIMENTER-EXAMPLE

"8 days"

EXPERIMENTER-EXAMPLE

"OK, the programme says -the duration Diane's headache is 8 days".

INSTRUCTION

3) "What does this piece of information tell you about the patient?"

INSTRUCTION

Answer the question, please...

INSTRUCTION

F) "Please attempt to predict what the next viable question or group of questions might be".

INSTRUCTION

Answer the question, please...

COMPUTER

4) How severe is Diane's headache (on a scale of 0 to 4 with 0 for very mild and 4 for very severe)

EXPERIMENTER-EXAMPLE

"How severe is Diane's headache (on a scale of 0 to 4 with 0 for very mild and 4 for very severe)"

INSTRUCTION

G) "I will this time select EXPLAIN, to give you and example of how to use this facility, and another window will be displayed on another part of the computer screen. You should see something like this in the EXPLANATION WINDOW."

---

E.G.EXPLAIN:

COMPUTER

Why are we asking about the severity of Diane's headache?

COMPUTER

Severity is a characteristic of headache that has not as yet been examined?
EXPERIMENTER-EXAMPLE

"Why are we asking about the severity of Diane's headache? The programme says it is because severity is a characteristic of headache that has not as yet been examined. This seems to be a rather general headache question".

INSTRUCTION

"At this point I will choose continue and the following will displayed: "

COMPUTER

**3

EXPERIMENTER-EXAMPLE

"3"

COMPUTER

5) Does Diane have a fever?

EXPERIMENTER-EXAMPLE

"Does Diane have a fever?"

INSTRUCTION

Around this point the experimenter will ask: "Please let us exchange seats so that you can operate the console. Do as I have done, so that you may become accustomed to the environment". Take this opportunity to ask any questions about what you are expected to do. The subject will go through the present case until they are comfortable (usually about 4 questions on and after using explanation once or twice).

INSTRUCTION

H) "Feel free to select, as often as you like and at any time, on the programme menu - 'explanation', before you 'continue'. This should explain the rationale for the current question in the CONSULTATION. Please read aloud what it says and think aloud as you make comment about the explanations".

INSTRUCTION

I) "If you have choosen 'explanation', say aloud if the explanation agreeable, or helpful? If the explanations are not, what alternative reasoning or action should be applied and why?"

INSTRUCTION

J) "Call up the next question at your own pace, and do the demonstration for as long as you like till you feel comfortable. Tell me when you want to run an actual trial which will be exactly as this, but with a different patient case"
K) "Please say aloud the answer and what it tells you about the patient". "Can you guess the next question(s)"? Can you make predictions about possible answers to the question(s) of the program and to your own?"

INSTRUCTION
L) "Continue, using EXPLAIN any time you desire it". ss

COMPUTER
6) What is Diane's temperature (in Fahrenheit)?

COMPUTER
**100. 2 Fahrenheit

page fills up

INSTRUCTION

M) When the window goes black push the space bar to go on,

-The previous windows material cannot longer be retrieved. Thank-you".

7) Please Elaborate

COMPUTER

7) Has Diane experienced recent seizures?

COMPUTER
**no

COMPUTER

8) Does Diane have vision problems?

COMPUTER
**Unknown

COMPUTER

9) Has Diane been exposed to any contagious diseases recently (e.g. meningococcal disease, mumps)?

COMPUTER
**no

COMPUTER

10) Does Diane have any cutaneous lesions or rashes on physical examination?
COMPUTER

**no

"Would you like me to go on, or do you know what to do now? If so then you have reached the end of the demonstration any questions? Now we will begin the experimental trial. Thank you for your cooperation".