HERAX:-
A MICROCOMPUTER-BASED EXPERT SYSTEM APPROACH FOR HUMAN RELIABILITY ANALYSIS

Abdelhamid ABDOUNI

Doctor of Philosophy

THE UNIVERSITY OF ASTON IN BIRMINGHAM

OCTOBER 1989

SUMMARY

HERAX: A Microcomputer-Based Expert System Approach for Human Error/Reliability Analysis

Abdelhamid ABDOUNI, PhD Thesis, University of Aston in Birmingham, 1989.

The groing awareness among managers and designers that human reliability is an important determinant of risk and profitability, and the need for tools to assist in the better understanding and systematic evaluation of human reliability with the view of incorporating the results within the overall framework of the probabilistic risk assessment, have lead to the development of various human reliability analysis (HRA) techniques.

Whereas acceptability and use of the conventional HRA techniques when carrying out global probabilistic safety and reliability studies is growing, many major obstacles have restricted their widespread application, such as the high level of knowledge, judgement and experience required to select and carry out a comprehensive HRA study, and the unavailability of human experts familiar with human factors and HRA techniques.

One of the main objectives of this research project was to investigate the feasibility of using Artificial Intelligence (AI) techniques to develop an expert computer program as one possible means of overcoming some of these obstacles.

The results of this study show that the characteristics of the HRA problem domain make it ideal for the expert systems technology, and that the expert developed here has potential benefits for the analysis of human reliability, such as ease of use, and reduction of the resources and time needed for the selection and application of the appropriate HRA technique.

The expert system approach proposed here, called HERAX, is intended to assist non-human factors specialists, such as designers and managers, in the selection and application of the appropriate HRA technique as part of a probabilistic risk analysis. With its explanation capabilities the system could also be used as a training tool.

The actual system is based on three of the existing HRA approaches, but could be easily extended to include more techniques. It also includes a list of the major factors that could affect human performance, as well as a data bank on human error probabilities.

Key words: Process Plant, Risk assessment, Safety, Ergonomics, Human Factors/Performance/Reliability, Errors, Expert Systems.

ACKNOWLEDGEMENTS

The author of this thesis would like to express his appreciation and deep gratitude to his supervisor Dr. Hani M. RAAFAT for his encouragement, valuable guidance and continued support throughout the various phases of this project.

The author also gratefully acknowledges the help and support of Professor Richard Booth, head of the Health and Safety Unit, Aston University, particularly for his assistance and suggestions regarding the presentation of the thesis given during the absence of Dr. H. Raafat.

The financial support that made this research work possible was provided first by the Ministry of High Education of the Algerian Government, and then by the Health and Safety Unit, Department of Mechanical and Production Engineering, Aston University, Birmingham, Great Britain.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Research Background

1.1.1 Changes in Process Control Plants

Within the past two decades modern systems such as nuclear power, chemical process, off-shore oil production and similarly systems have undergone considerable changes (see Chapter 3). Process conditions such as pressure and temperature have become more severe. The concentration of stored energy has increased. Plants have grown in size and are becoming increasingly complex, costly and high risk as a result of the high automation and centralization in order to secure effective and economically optimal operation.

1.1.2 Implications of Process Changes

These factors have greatly increased the potential for large-scale accident hazards such as fires, explosions and toxic releases which can have serious consequences in terms of loss in human life, damage to the environment, and economic costs.

Major disasters, such as Flixborough (1974), TMI (1979), Bhopal (1984), and Chernobyl (1986), have increased public and governmental concern and the adequacy of existing process safety methods has been subject to an increasingly critical examination. This has led to the development of new assessment techniques. These approaches differ somewhat from safety analysis as traditionally conceived in the process industries. The main difference is the increased use of systematic, quantitative and probabilistically oriented analyses over more traditional engineering judgement (see Chapter 2).

1.1.3 Safety Improvement and PRA

Probabilistic/Quantitative risk (or safety) assessment (PRA, QRA, or PSA) technology is being used and has proven to be an important method for improving the safety of plant design and operation practices.

PRA has not yet become a formal requirement of the licensing or regulation of hazardous or nuclear installations (see Chapter 4). However, since the US Nuclear Regulatory Commission's publication of the Reactor Safety Study (NUREG-75/014, WASH-1400) in 1975, the use by the nuclear power industry of PRA has increased dramatically. Approximately 30 PRA studies have been or are being performed in the US alone and more than that figure have been completed so far for nuclear power plants in various countries with the assistance of the International Atomic Energy Agency (IAEA). *"These studies have yielded plant-specific, safety-related insights which could not have been obtained by any other means."* (Lederman and Gubler, 1987).

A number of companies in industries other than the nuclear industry have also begun to make significant use of probabilistic risk assessment in the last decade, as part of their decision making process on design safety issues (for example, Canvey Island in the UK, and the Rijnmond area in The Netherlands).

1.1.4 Process Automation and Human Operator's Role

High levels of automation has been introduced and various countermeasures have been implemented in the design, construction, operation and maintenance stages to ensure safety and reliability of modern process control plants. However, the Chernobyl and TMI accidents have shown that despite the introduction of automation and the

implementation of necessary mechanical improvements, human errors are still happening and can lead to serious accidents.

Because automation can fail to work properly and because designers of automatic systems cannot anticipate every plausible response need of a man-machine system, the role of human operators is still important. The human operator continues to play a key role in all aspects of plant operation, including design, maintenance, testing, and management of these plants. Furthermore, there arise situations where the operator's skills such as diagnosis and decision making would be required, particularly in cases of unforeseen problems, e.g., start-up, shut-down, automatic system failure and emergency procedures. A way of providing a balance between automation and human actions has, therefore, to be found, (see Chapter 9, for more discussion).

1.1.5 Human Performance and Plant Safety/Reliability

There is a body of evidence to support the notion that humans play a dominant part in contributing to major accidents at industrial facilities (see Chapter 5).

Contributions of human actions have been found both beneficial and detrimental. Detrimental actions include, for example, those that result in the unavailability of plant systems before an initiating event, and those which cause an initiating event to occur. Beneficial actions include correct diagnosis of the nature of an event and the implementation of recovery procedures.

Accidents caused by human operators are still happening and have not decreased and accident statistics and analyses clearly indicate the necessity to take into account and consider more explicitly the contribution of human reliability to the overall system safety, reliability and availability.

It is known, for example, that 70% of aviation accidents are due to crew error (Flight International, 1975); and empirical and analytical studies show that,

> *"human error contributes about one half to the risk of accidents in nuclear power plant operation"*, (Blix, 1988)

Some of the recent major accidents, e.g., Flixborough, TMI and the incident at Chernobyl occurred as a direct result of human error.

More recently, a discussion panel at a conference on *"Man-Machine Interface in the Nuclear Industry"*, organized by the International Atomic Energy Agency (IAEA), has concluded that:

> *"The ultimate line of defence against accidents is man, with his capability for flexibility and innovative thinking."* (IAEA, 1988).

It was also mentioned that PRA is a useful tool to analyse the vulnerability of a plant.

1.1.6 Plant Safety/Reliability and Need for HRA Techniques

In view of the importance of human reliability as a major determinant of system safety and reliability, it is, therefore, clear that in order to enhance system safety and reliability, and reduce the potential for disasters, human reliability needs to be improved and the probability of human error needs to be reduced. This objective could be achieved by:

1. Applying current knowledge from behavioural sciences to develop procedures, or "*Human Reliability Analysis*" techniques (HRA), which systematically consider human factors during plant safety and risk assessment analysis, from both the qualitative and quantitative stand-points.

2. Making the knowledge and experience of Human Factors/Ergonomics experts in human reliability analysis available to those who need it, such as engineers and managers.

1.2 Related Work

During the project development procedure, a background research was conducted to determine what work has, or is being done in the problem area. The main results of this review are briefly described here (see Chapter 2 for more details).

1.2.1 Computer Application to Safety-related Problems

There are many ways in which risk assessment techniques in general and human reliability analysis techniques in particular can be improved, and hence, safety and reliability of process plants enhanced. One promising way is the use of computers (Andow, 1987). In the last past two decades, computer technology has increasingly being used in safety and reliability domains (see Kibblewhite, 1988 for a detailed review of safety-related computer applications).

Until recently, the safety and reliability problems solved by conventional computer techniques are so-called well-structured problems where procedures for problem-solving are completely predetermined. However, there are many ill-structured safety-related problems, including human reliability analysis, that require human judgement and experience for their solutions. In solving these types of problems by computers, it is necessary to encode the knowledge and the methods of inference the human experts have. Conventional computer programs are not suitable for these types of problems.

1.2.2 Artificial Intelligence and Expert System Technologies

Much attention has been paid to the application of the subfield of computer science called "*Artificial Intelligence*" (AI), and in particular to its subdiscipline "*knowledge-based expert systems*" technology. Expert systems have many advantages over traditional computer programs (see Chapter 2), and are increasingly seen as a new and powerful tool for assisting management and process designers in their problem-solving and decision-making tasks. The application of expert system technology to risk assessment, and in particular to human reliability analysis, could contribute significantly to the improvement of plant safety and reliability.

Expert systems have a very broad range of application (see Chapter 2), but are particularly useful where:

- The problem domain is complex and involves the choice of one or more possible alternative decisions.
- There is little underlying theory but rather a general body of knowledge.
- There is a shortage of experts.

1.2.3 Expert Systems Application to Safety-related Problems

Problem solving in the fields of risk and safety management has many of the characteristics well-suited for expert system applications. Often several alternative solutions are available; and, in general, there is a body of knowledge or set of standards rather than a comprehensive underlying theory on which decisions are made. Also, there is a shortage of experts and the data used in decisions are noisy (Baybutt, 1985).

The use of AI and expert system technology for safety applications covers a large range of areas (see for example, ANS Topical Meeting, 1985; Colley, 1985, IBC, 1988). Applications of expert systems in the fields of safety and reliability include their use in (see Chapter 2):

- Hazard identification and assessment
- Decision making
- System safety and reliability analysis
- Process control
- Alarm handling (reduction)
- Diagnosis of system faults
- Inspection and maintenance.
- Operator training

Most of the expert systems for safety and reliability analysis, described in the literature, are systems developed or under development as prototypes, but have not been completely developed and applied in plants.

Most of the previously published work on application of expert system technology to safety-related problems was restricted to hazard identification and analysis techniques, such as HAZOP (Hazard and Operability Studies), system (hardware) reliability analysis techniques, such as fault trees an FMEA (failure modes and effects analysis), and human performance aids. Nothing had, however, as far as this author is aware, been done on the quantitative analysis of human reliability in nuclear and process industries.

1.2.4 Characteristics of HRA Knowledge

The domain literature reviews carried out during this project show that despite being a critical component of any risk/safety study, the widespread application of HRA techniques by non-experts in human factors/ergonomics has been limited. Some of the main deficiencies that have characterized HRA knowledge and have limited their use are as follows:

(i) The growing number of analytical techniques developed. Several *HRA* techniques have been developed as a means to reduce human errors and improve plant safety/reliability, particularly in the nuclear industries, and continue to be refined. Three major categories of HRA approaches have been identified (see Chapter 6). These are:

1. Analytical Approaches.- Based on task analysis.
2. Time-dependent Approaches.- Based on classical reliability theory.
3. Subjective-judgement Approaches.- Make much use of quantified expert judgement.

(ii) Considerable variability was observed between the human reliability studies (and to some degree within a given study) as to the selection and application of an approach to quantify human errors.

(iii) Human behaviour is a complex subject that does not lend itself to simple models like those for component and system reliability. This makes the analysis of problems involving human actions more dependent on judgement, experience and insights of Human Factors/Ergonomics experts who are familiar with a wide range of such problems. They are capable of reasoning about new problems based on their

experience, in-depth knowledge and advanced problem-solving techniques which very often involve "rules of thumb".

(iv) The application of human reliability analysis procedures involves a series of complicated analytical steps not always clear, and, in addition to being complex and time consuming, it requires specialized training and thorough knowledge in many disciplines, such as psychology, engineering, and statistics.

(v) The level of training and experience required to build this expertise is high and takes many months of apprenticeship with a more experienced human reliability expert. Only a handful of such experts exists and are in great demand.

In the light of the above comments, it is clear that the various existing HRA techniques are complex and rely heavily upon human factors/ergonomics analysts' judgment and expertise for selecting and applying the appropriate models and particularly for selecting data for quantification. Such complexities make the selection and conduction of these techniques by non-specialists in human factors very difficult.

These characteristics are some of the main obstacles to the widespread application of HRA techniques by non-human factors/ergonomics specialists, such as engineers, designers, managers, and all those concerned with safety and reliability of process control plants, when carrying out overall system safety and reliability assessments.

There is thus, a need for tools that could provide some guidance to non-experts in human factors in the selection and use of the appropriate HRA technique. It is also clear that being knowledge-intensive and experience-based make HRA problems suitable for expert system technology application. Therefore, the development of an expert system is one possible means to overcome some of the above-cited HRA shortcomings and would, consequently help their widespread use by non-experts.

1.3 Research Problem Overview

Some of the questions that this research project attempts to address are as follows:

a) What is expert system technology?

b) What HRA techniques exist already, and what is their state of the art development?

c) Is there any single fully developed HRA technique to be suitable candidate for the proposed expert system? If there is not,

d) Can more than one of the available techniques be combined and used in a new format to be implemented in the proposed expert system?

e) What are the requirements and incentives for the development of the expert system for HRA?

1.4 Research Objectives

The main objectives of this research project are to:

1. Investigate the feasibility of using AI techniques to develop an expert system approach for the modelling and quantitative analysis of human reliability as part of a probabilistic risk assessment (PRA) in nuclear and major hazard process plants.

2. Develop and refine an experimental prototype expert system to this end with the following characteristics:

(a) provide a systematic analysis framework to aid the non-human factors experts in the selection and systematic conduction of the appropriate

human reliability analysis technique as part of overall PRA studies for their specific process plant under normal as well as abnormal situations.

(b) include some of the existing HRA approaches.

(c) include a data bank on human error probabilities.

(d) reduce the resources and time needed for HRA conduction.

The expert system for HRA proposed by this research project and characterized by the above features is expected to contribute to the broad usage of HRA techniques and could be considered as a complementary and worth doing effort to the work already done in the application of expert systems technology in other fields of hazard/risk and safety management.

1.5 Research Scope

To meet the above research objectives, an expert system for HRA, called **HERAX** (**H**uman **E**rror/**R**eliability **A**nalysis e**X**pert), has been developed. This program is written in Common LISP and runs on IBM-PC AT and IBM compatibles under the operating system DOS.

Numerous techniques for the modelling and quantification of human reliability have been reported in the literature. From these, three have been selected to form the knowledge base of the expert system, namely,

a) THERP: Technique for Human Error Rate Prediction (Swain et al., 1983),

b) SLIM: Success Likelihood Index Method (Embrey et al., 1984), and

c) APJ: Absolute Probabilities Judgement method (Comer et al., 1984).

Although, the current number of HRA approaches included in the system is limited to three, it can easily be extended to include more techniques.

While more limited in detail than the full versions of the HRA techniques on which it is based, this new expert system HRA approach is far more accessible, easy to use and update.

The expert system approach developed here is intended to be used for the *modelling* (identification of human error or qualitative analysis) and *quantitative analysis* of human performance in nuclear and major hazard process industries.

The system is mainly designed to be used as part of a probabilistic risk analysis study (PRA) for nuclear and process plants. However, it may also be used as a design or teaching tool.

The main target of this expert system is safety and reliability in nuclear and major hazard process industries. However, the program can be applied (in its current form, or if not applicable, modified and adapted) to other industrial and occupational situations.

It is emphasized that the system, at this stage, is only an aid to the less experienced analysts in the selection and application of three HRA procedures, and therefore must not be viewed in any sense as a potential replacement for the HRA expert.

1.6 Research Strategy

The methodology employed in carrying out this research included the following tasks:

(i) Acquisition of the knowledge base necessary to develop a prototype expert system for HRA by carrying out:

a) a critical review and analysis of Artificial Intelligence, and particularly Expert System techniques, with respect to their benefits, limitations and actual applications to HRA-related problems.

b) a critical review and analysis of conventional approaches to HRA in process control plants in the context of their completeness, ease of use, and suitability to expert system technology.

(ii) Selection and combination of three current HRA techniques into a new interactive and easy to use format.

(iii) Familiarization with the computer (IBM-PC/AT), and learning of the computer programming language (LISP) used for the implementation of the proposed expert system.

(iv) Development and review of an experimental prototype system and refinement of the HRA techniques selected taking advantage of the capabilities of the computer and flexibility of expert systems technology.

(v) Validation of the system. Having established the feasibility of using expert system techniques to HRA, collected the basic knowledge, and developed a prototype system, the larger question of its effectiveness warranted review. To this end, many attempts have been made to get a financial support from industrial companies for the application of the system to real cases. However, all these attempts have failed Therefore, some case studies from the literature were selected and used in which different abnormal situations in chemical process industries were analysed using the expert system approach developed during this research project (see for example, Raafat and Abdouni, 1987).

1.7 Thesis Structure

The thesis is divided into the following chapters:

Chapter 1, *"Introduction"*, this chapter, in which the background to this research project has been discussed and a summary of related work has been provided, followed by a discussion of the research problem, scope, strategy and thesis structure.

Chapter 2, *"Review of Related Work"*, starts with a definition of artificial intelligence and expert systems respectively. This is followed by a description of the basic expert system components. A discussion of the ways expert systems differ from conventional computer programs is then presented. The tasks involved in building expert systems are described. Finally, applications of expert systems in different fields are reviewed with more emphasis on safety and reliability domains.

Chapter 3, *"Safety in Nuclear and Major Hazard Industries"*, presents a historical overview of safety in nuclear and major hazard industries. It then describes the probabilistic risk or safety assessment techniques followed by a discussion of their application in both industries. Finally, conclusions about their usefulness and limitations are given.

Chapter 4, *"The UK Approaches to Risk Assessment and HRA"*, examines the UK approaches to risk quantification and human reliability analysis in the nuclear industries and major hazard industries, both from the point of view of the regulatory requirements for quantification of risk and the quantification practices that are adopted in reality.

Chapter 5, *"Qualitative Approaches to Human Reliability Analysis"*, presents a detailed theoretical analysis of the expert system problem-domain, i.e., the main conceptual

background underlying the areas of human reliability and human errors. It includes a historical review of both ergonomics/human factors, and human reliability, definitions of human reliability and human error, modelling of human performance, and finally, classification of human errors.

Chapter 6, *"Conventional Quantitative Approaches to Human Reliability Analysis"*, gives an account of the review carried out to identify the existing HRA techniques. It includes a discussion of the problem of human performance data, followed by a brief review of the major classes of HRA techniques available.

Chapter 7, *"Selected HRA Methods"*, provides a detailed description and discussion of the three techniques selected for inclusion in the expert system for HRA, namely: THERP, SLIM, and APJ.

Chapter 8, *"Development of HERAX: Human Error and Reliability Analysis Expert System"*, describes the development process, structure and main functions of the expert system approach which has been developed during this research project.

Chapter 9, *"Application of HERAX"*, presents a case study that is used here to demonstrate the application of the expert system approach for HRA to the analysis of two typical abnormal situations in a process plant, both involving human operators interventions.

Chapter 10, *"Conclusions and Recommendations for Future Work"*, summarizes the findings of the research, and presents some recommendations for potential improvements to the system developed.

CHAPTER 2

REVIEW OF RELATED WORK

2.1 Introduction

Since their inception in the late seventies, expert systems have rapidly propagated and developed in various fields of science, engineering, medicine, and business. One major area in which several expert systems have been developed is medical diagnosis. MYCIN: bacteriological blood infection diagnosis (Shortliffe, 1976), and INTERNIST-1: diagnostic assistant for general internal medicine (Miller, et al., 1982) are among the best known systems. Expert systems have also been developed for other applications. Examples include Xcon (or R1), for designing and configuring computer systems (Mc Dermot, 1982); PROSPECTOR (Duda et al., 1979) and Drilling Advisor, for assisting with the evaluation of mineral and drilling data and problems; DENDRAL, for the solution of complex molecular design problems (Buchanan, and Feigenbaum, 1978; Lindsay, et al., 1980); and SOPHIE, for the diagnosis of electronic and mechanical malfunctions (Brown et al., 1982).

Most of these systems have been developed and revised over a number of years. Although not many of these systems have proven successful when actually used, some of them have gained commercial acceptance; among these is Xcon (or R1).

The interest in expert systems application to process plants safety and reliability has also increased, particularly in nuclear power plants, as evidenced by the number of papers on AI applications presented at the *"First International American Nuclear Society Topical meeting on Computer Applications for Nuclear Power Plant Operation and Control"* (ANS, 1985), and more recently, at the conference on *"Expert Systems and Industrial Hazards"* (IBC, 1988), held in London, where several reports on expert

systems developed or under development for process safety have been presented. These include: "safety integrity management", "Hazard evaluation", and "Rapid fault diagnosis".

The purpose of this chapter is, first to briefly introduce a survey of expert system techniques (for a comprehensive state-of the art review, see, for example, Hayes-Roth et al., 1983; Alty and Coombs, 1984; Weiss and Kulikowsky, 1984; Harmon and King, 1985; Waterman, 1986). It, then reviews their potential and current applications to safety and reliability problems.

2.2 Artificial Intelligence and Expert Systems

2.2.1 AI Definition

There is no agreed upon definition of the subfield of computer science known as Artificial Intelligence (AI). According to Trappl (1986), the definition of AI most widely accepted (particularly in the United States) is *"making computers smart"*, i.e., intelligent (behaviour oriented approach). Another definition, also widely accepted is *"making computer models of human intelligence"* (cognitive approach).

The Handbook of Artificial Intelligence by Barr and Feigenbaum (1981-82), states:

> *"Artificial Intelligence is the part of computer science concerned with designing intelligent computer systems, that is, systems that exhibit the characteristics we associate with intelligence in human behaviour - understanding language, learning, reasoning, solving problems, and so on."*

This definition concentrates on the comparison between the abilities of humans and the abilities of computers. The following definition of AI focus on the difference between

programming techniques used in AI, particularly in expert systems, and more conventional methods of programming.

Buchanan and Shortliffe (1984), state that:

> *"Artificial intelligence is that branch of computer science dealing with symbolic, nonalgorithmic methods of problem solving."*

Another important AI concept is *heuristics*. AI researchers rely on heuristics (rules of thumb) to solve problems. They also use *pattern matching* techniques in an attempt to discover relationships between objects, events, or activities.

2.2.2 AI Evolution and Expert Systems

AI researchers have always been interested in making computers "intelligent" or "think". Early in the history of AI, many researchers believed that by simulating the complicated process of human reasoning or thinking, computers could solve problems without having access to large amounts of specific knowledge. Although early attempts to solve problems with pure reason seemed promising, they ultimately proved to be unsuccessful.

Many attempts were made to develop general methods for problem-solving, i.e., programs were designed to address several problem areas by applying some general problem-solving strategy. Although some interesting programs were written, there were no real breakthroughs.

The next phase of AI research, in the early 1970's, narrowed down the problem domain to a single problem and concentrated on developing techniques to formulate a problem; this was known as "representation". This was combined with developing techniques to control the search for a solution such that it would not take too long or

consume too much computer memory. Again some very interesting work was undertaken without any real breakthrough in terms of making a computer "think".

The real breakthrough came in the late 1970's with the realization that the problem-solving power of a program comes from the knowledge it possesses and not just from the programming techniques and control schemes or formalisms it employs (Feigenbaum, 1977). This AI goal has been stated quite simply by Waterman (1985), as follows:-

> *"To make a program intelligent, provide it with lots of high quality, specific knowledge about some problem area."*

This realization led to the development of special-purpose computer programs that reason with knowledge, and that were "expert" in a narrow domain. These programs were called "Expert Systems" or "Knowledge-based Expert Systems".

Expert systems technology is currently the most well-known and one of the most successful branch of AI. Other research areas of the AI field include (see Figure 2.1):

- problem solving,
- logical reasoning,
- natural language processing,
- automatic programming,
- robotics,
- speech recognition and synthesis,
- computer vision,
- learning,
- planning and decision support.

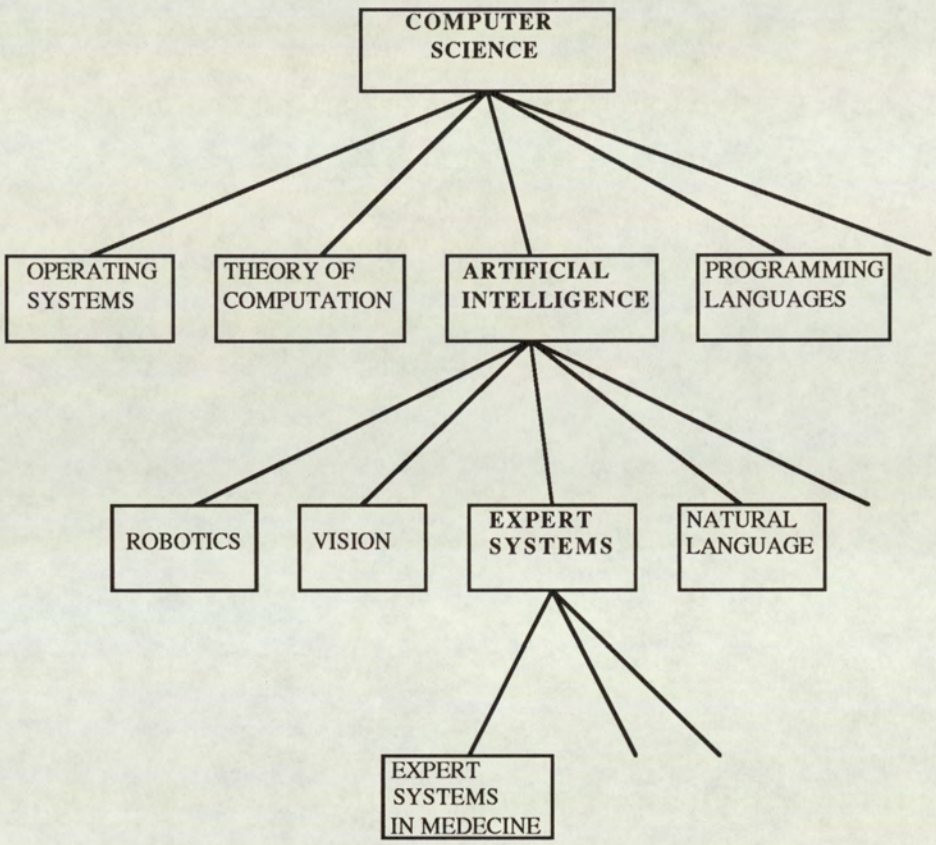


Figure 2.1 Computer Science, Artificial Intelligence, and Expert systems

The underlying theme in the AI work, and in the wide range of disciplines it spans, such as psychology, linguistic, philosophy, mathematics, and cognitive sciences, is the emphasis on symbolic computation, and the development of heuristic solutions to complex problems.

2.2.3 Expert Systems Definition

There have been many definitions of "expert systems", but at a general level they are computer programs that attempt to emulate experts by capturing and representing knowledge, expertise or problem-solving capabilities in specific well-defined subject that are perishable, scarce, vague, and difficult to apply, distribute, or accumulate, and make that expertise and advice available to those who need it. This "built-in" knowledge enable expert systems to provide expert-level performance where human experts are either unavailable or not cost effective. When key personnel retire, transfer, or quit, expert systems can smooth the transition and serve as a training tool for new personnel.

Feigenbaum (1982), describes an expert system as follows:

> *"An expert system is an intelligent program that uses knowledge and inference procedures to solve problems that are difficult enough to require significant human expertise for their solution. The knowledge necessary to perform at such a level, plus the inference procedures used, can be thought of as a model of the expertise of the best practitioners of the field."*

2.3 Overview of Expert System Technology

2.3.1 Expert System Components

An expert system can contain from three to six components. All expert systems share the following common fundamental architecture, regardless of how they are implemented.

1. A knowledge base.
2. An inference engine.
3. A working memory.
4. A user interface.

The basic structure of an expert system is shown schematically in Figure 2.2.

1. Knowledge Base

A knowledge base contains both declarative knowledge (facts about objects, events, and situations) and procedural knowledge (rules that state relations between facts of the problem domain or information about courses of action). The inference engine, can sort through the rules in the knowledge base to analyse the problem and conclude what action or actions to take.

Although many knowledge representation tools and techniques have been used in expert systems (see for example, Jackson, 1986), the most prevalent form of knowledge representation currently used in expert systems is *"the rule-based production system"* approach.

(a) Production Rules

Rule-based systems, especially production rules constitute the best currently available means for codifying the problem-solving know-how (rules of thumb) of human experts (Hayes-Roth, 1985). They excel in flexibility, modularity and expandability.

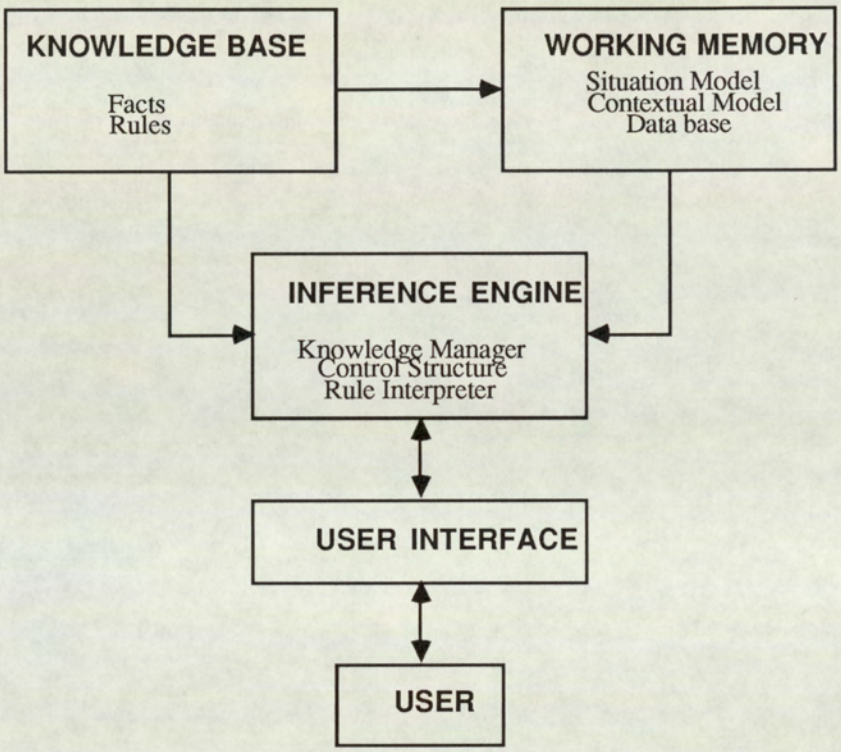


Figure 2.1 Expert System Components

Within the production system formalism, it is possible to express different levels of knowledge as well as different problem-solving strategies.

Each rule in a production system consists of two parts. The left-hand side or antecedent (premise) describes some pattern or situation that must be matched. The antecedent typically contains several clauses linked by AND's and OR's. The right-hand side or consequent (conclusion) describes an action or actions to be taken or information to be gained as a result of employing, (activating, firing or executing) the rule. These new facts can themselves be used to form matches with the "IF" portion of other rules and so chains of inference can be produced. Most often, this sort of rule is referred to as an "IF/THEN" rule or simply a *"production"* and expressed in the following conditional form:

IF (premise) FACT1, FACT2,...
THEN (conclusion) FACT10, FACT11,...

Another method commonly used for representing knowledge in expert systems is *"frames"*.

(b) *Frames*

Frames (also called *units*) are discrete structures having individual properties (*slots*) into which all domain knowledge is partitioned. Frames can be used to represent broad concepts, classes of objects, or individual instances or components of objects. A frame associates *attributes* (which can be assigned or inherited from other frames) with objects. These attributes are filled in the slots of a frame.

2. Inference Engine

The inference engine is sometimes referred to as the *"knowledge manager"*, the *"control structure"* or the *"rule interpreter"*. This is the main controlling component. It has the task of deciding which rule or rules to apply and when. There are two important ways in which rules can be used in a rule-based system; one is called *"forward chaining"* and the other *"backward chaining"*.

a) Forward Chaining

A problem solver is doing forward chaining if it starts with a collection of facts about the situation considered and searches for a matching rule, trying all available rules over and over, adding new facts (concluded as a result of rules' execution or firing) to the knowledge base as it goes, until no rule applies. The forward chaining approach looks for rules that depend only on known facts. This is why forward chaining is sometimes associated with the term *"data driven"*

b) Backward Chaining

Backward chaining on the other hand, starts with the hypothesis, goal, or "action" clause of a rule and searches the knowledge base for facts which prove the hypothesis (conclusion), i.e., determine whether the "condition" clause (premise) of the rule matches the situation. This is sometimes termed *"goal driven"*.

Which method is best depends upon the nature of the problem. However, both chaining techniques can be combined in an expert system. Initially, forward chaining is used and volunteered data is accepted from the user. After this stage the system switches to backward chaining and asks for facts to be confirmed. In the forward chaining mode the system is abducting hypotheses from the initial input data. In the

backward chaining mode the system is seeking to confirm these hypotheses by asking the user about the antecedents of the rules having the hypothesis as a consequent.

3. Working Memory

This is a workspace area, or situation model (facts or assertions about the current problem being solved) to represent the user problem and provide scratch pad space to work out the problem solution. In normal operation the inference engine employs the information contained in the knowledge base to interpret the current contextual data in the situation model.

4. User Interface

The user interface module is used to communicate with the user. The communication performed by a user interface is bidirectional. The user is asked to describe the specific problem under consideration and the computer system generates conclusions. If requested, the system will explain *why* certain input information is required and *how* it reached a particular conclusion (i.e., which rules were used and when).

A good user-friendly interface to the system will assist considerably its acceptance and subsequent use.

2.3.2 Expert Systems Vs Conventional Programs

Expert systems differ from traditional, data-based software programs. How they differ? One way of answering this question is by pondering a related question: *"What makes a human being an expert?"* One possible response is that the person must be able to solve problems at a high level of performance. But problem-solving

performance alone is not sufficient basis for calling someone an expert, much more is expected from a human expert.

Similarly, a standard operations research algorithm would not qualify as an expert system, even though it can easily surpass a human in computational performance.

Key attributes of what experts are capable of can be summarized as follows:

- they solve problems
- explain their results
- learn any thing new about the domain
- restructure their knowledge
- break rules
- determine the relevance of their knowledge
- deal with missing or imprecise information

In more pragmatic terms, expert systems differ in a number of ways from conventional computer programs. First of all, in an expert system, the control structure (decisions, choices, conclusions, results) is separated from the domain-specific knowledge. This is different than in programs written in FORTRAN or other "procedural" languages, with which domain changes require the entire program to essentially be rewritten, including control structure or statements. Thus, traditional programs are very inflexible.

A major difference between expert system and traditional data processing programs is the sequence. Conventional programs tend to be sequential in execution. In an expert system, the rules and facts in the knowledge base can essentially be in any order, and the inference engine automatically applies the correct knowledge when certain patterns

exist in the data. This is possible because expert systems manipulate symbols rather than numbers.

Most current expert systems display some of the human-like qualities that make dealing with human experts so appealing, namely coping with uncertainty in data and knowledge, explaining how an answer was arrived at and modifying the current knowledge base when something new is learned. Another advantage of expert systems is the potential for achieving completeness. That is, where an expert's knowledge is insufficient, it can be augmented by that of another expert.

One final difference between expert systems and conventional approaches is that a traditional computer program can be viewed as a series of steps which tell the computer *"How to do"* a particular task, whereas the method used by expert systems is geared to *"What to do"*.

2.3.3 Desirable Qualities of an Expert System

An expert system, to perform intelligently and efficiently, should possess certain characteristics. Although each expert system has its own particular characteristics, there are several features common to many systems. The following list (Buchanan and Shortliffe, 1984), suggests seven criteria that are important prerequisites for the acceptance of an expert system by its intended users. These criteria form a useful list of features that are desirable in any expert system. The program should be:

1. Useful.- An expert system should be developed to meet a specific need, one for which it is recognized that assistance is needed.
2. Usable.- An expert system should be designed so that even a novice computer user finds it easy to use.

3. Educational when appropriate.- An expert system may be used by non-experts, who should be able to increase their own expertise by using the system.
4. Able to explain its advice.- An expert system should be able to explain the "reasoning" process that led it to its conclusions, to allow the user to decide whether to accept the system's recommendations.
5. Able to respond to simple questions.- Because people with different levels of knowledge may use the system, an expert system should be able to answer questions about point that may not be clear to all users.
6. Able to learn new knowledge.- Not only should an expert system be able to respond to the user's questions, it also should be able to ask questions to gain additional information.
7. Include knowledge easily modified.- It is important that the user be able to revise the knowledge base of an expert system easily to correct errors or add new information.

Certainly, not every expert system built to date possesses all of the features mentioned. Each feature may be viewed, instead, as constituting a continuum ranging from none to an ideal amount of the features

2.3.4 Expert System Development

Any discussion of developing an expert system must include the knowledge engineering process and a description of the tasks of the individuals who do knowledge engineering.

(i) Knowledge Engineering

The terms *knowledge engineering* and *knowledge engineer* were first coined by Feigenbaum and the researchers at Stanford University. The term "knowledge

engineering" is used to describe the process one goes through to develop an expert system. The term knowledge engineer is used to describe the person who actually develops such a system.

The tasks of a knowledge engineer can be described in terms of three subtasks (Harmon et al., 1988):

(a) Knowledge acquisition.

(b) Knowledge modelling.

(c) Knowledge encoding.

(a) Knowledge Acquisition. Knowledge acquisition (or knowledge elicitation) includes all activities involved in obtaining information from human experts or other sources.

(b) Knowledge Modelling. Knowledge modelling or representation is the process by which knowledge engineers organize the information they acquire from the experts.

(c) Knowledge Encoding. Knowledge encoding is the process of actually entering facts, rules, objects, and relationships information into an expert system.

The knowledge engineer requires a mixture of psychology, adaptability and common sense (Gallacher, 1989).

(ii) *Expert System Building Methodology*

Expert system development can be viewed as five highly interdependent and overlapping phases (Hayes-Roth et al., 1983).

1. Identification (definition).
2. Conceptualisation.
3. Formalization.
4. Implementation.
5. Testing.

Phases 1 and 2 are concerned with the nature of the problems that the expert system is intended to solve and the vocabulary it uses in their solution. Phase 3 is concerned with the design of the system architecture and knowledge representation. Phase 4 is concerned with the transfer of the knowledge base to the computer.

Phase 1: **Identification**

This activity involves knowledge acquisition and familiarisation with the problem domain. The knowledge engineer will learn as much as possible about the problem domain by interviewing one or more human experts or by reviewing the literature, both theoretical and specific. This will enable him to determine the important features of the problem and how the expert system is expected to contribute to its solution. It is recognized to be one of the main obstacle to successful expert system development (Buchanan et al., 1976; Shortliffe, 1976; Alty et al., 1985). It can be a long and painful one requiring repeated trial and correction.

Phase 2: **Conceptualisation**

During the conceptualisation stage, the knowledge engineer frequently designs a diagram of the problem as an aid to the decision making concerning the concepts, relations and control mechanisms that are needed to describe problem solving in the domain. This process usually involves dividing the problem into a series of

subproblems and drawing both the relationships among the pieces of each subproblem and the relationships among the various subproblems.

Phase 3: **Formalization**

Whereas the focus during the preceding stages was on understanding the problem and proposing a solution, the effort here is to relate the domain problem to the expert system technology that may solve it. Formalization involves selecting a development technique, a strategy, or a format appropriate for organizing and representing the knowledge and expressing the key concepts and relations in some formal way.

Formalization is a critical part of the development process, requiring great skill on the part of the knowledge engineer. Many domain experts describe *what* they do but not *why*; therefore, one of the knowledge engineer's primary responsibilities is to analyse example situations and distil from those examples a set of rules that describe the domain expert's knowledge. This process is the most time consuming stage.

Phase 4 **Implementation**

During implementaion the formalized knowledge is programmed into a the computer that has been chosen for system development, using the predetermined techniques and tools. A first cut at creating a knowledge base can be made and a prototype of the expert system produced.

If the first prototype works at all, the knowledge engineer may be able to determine if the techniques chosen to implement the expert system were the appropriate ones. On the other hand, the knowledge engineer may discover that the chosen techniques simply cannot be implemented. At this point, the concepts may have to be

reformalised, or it even may be necessary to design new development tools to implement the system efficiently.

Phase 5: **Testing**

Testing involves evaluating the performance and usefulness of the prototype system and revising it if necessary. Results from the tests are used as "feedback" to return to a previous stage and adjust the performance of the system.

The expert system building stages discussed above are not clear-cut, well-defined, or even independent.

Although great strides have been made in expediting the process of developing an expert system, it often remains an extremely time-consuming and non-trivial task. The development of a sophisticated system may require a team of several people working together for more than a year.

Since the problem domain may change or become better understood over time, an expert system is never really complete. An expert system typically is developed and refined over a period of several years. It is important, therefore, that a system be flexible enough to withstand many modifications.

2.3.5 E.S Building Vs Conventional Program Building

The process involved in constructing an expert system (ES) is not much different from other software development efforts (see Figure 2.3). However, there are a few key differences. One of these is that knowledge base expert systems developed early in the process can be discarded or completely restructured.

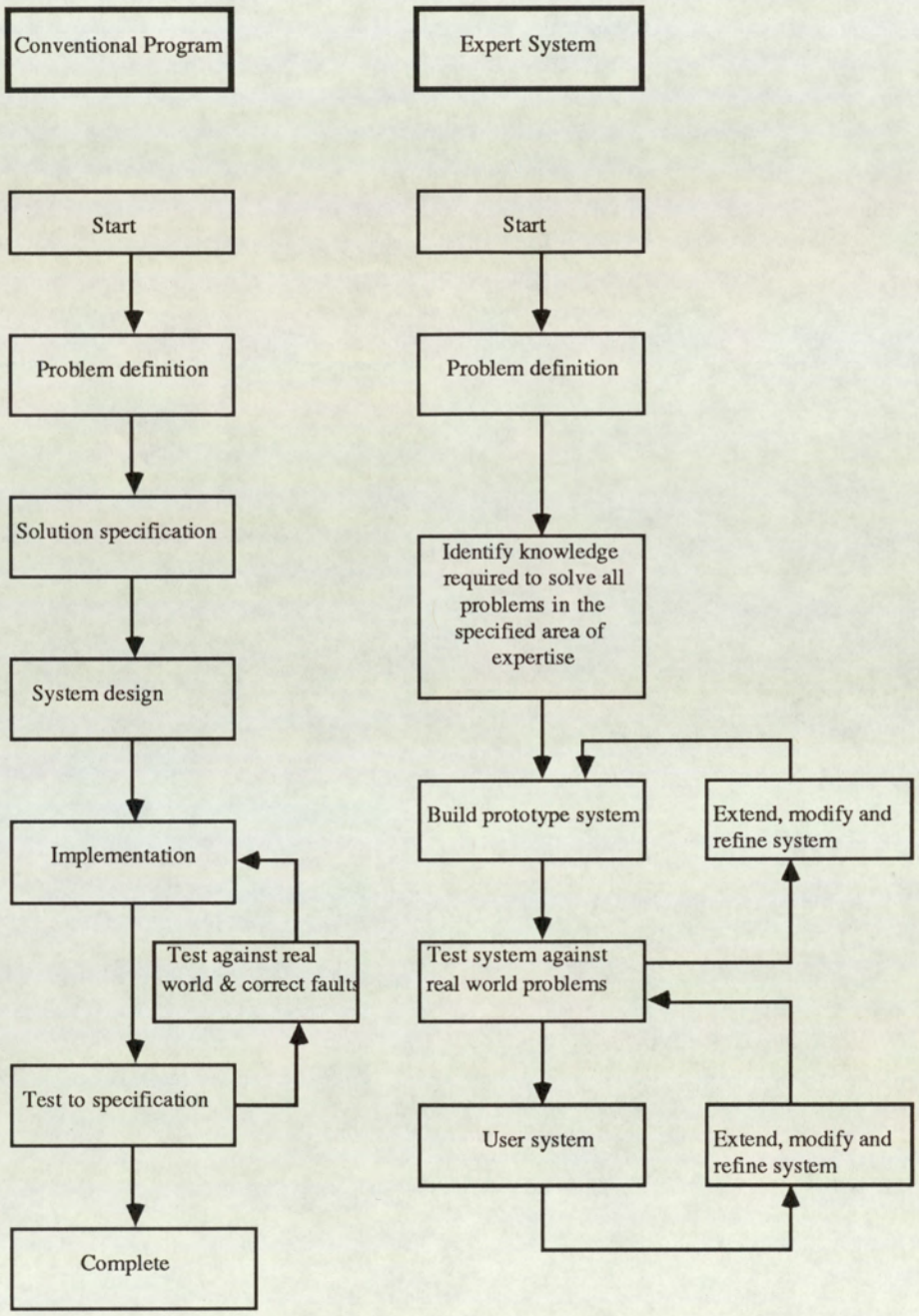


Figure 2.3 Differences between Conventional Program Development and Expert System Development

In contrast, conventional programs are developed sequentially. Contractors or end-users decide what they want and communicate their requirements to software engineers. The software engineers design an overall system to meet the contractor's requirements. They partition the system into small, but interdependent, modules that many individuals can work on. Then they check out the system requirements and design with the contractor, freeze those requirements and design, and implement, test, and deliver the system.

Clearly, if the contractors do not know what they want, the program design will be incorrect. If the software engineers misunderstand the requirements that the contractors intended, their completed system will be unacceptable. If, after the system is built, the builders and contractors discover important new features that make the system far more suitable, they must tear down the system and rebuild it at great expense. A change in even one part of a program generally necessitates a rebuilding effort, because in conventional programming a change in one place is likely to impact the rest of the program.

In contrast, expert system techniques allow developers to explore ideas and to change their minds. This is made possible by the separation of the program structure and mechanisms that contain knowledge, infer information, and control the way the program infers information. This separation makes it relatively easy to modify one part of an expert system and to add new knowledge incrementally, without affecting other parts of the program. This is a major advantage over conventional programs.

2.3.6 Expert System Building Tools

The tools that may be used in building an expert system fall into two broad categories. Firstly, a general purpose programming language can be used. When this is the case

the expert system is tailor made. Secondly, an expert system may be built using a "Shell" system.

(i) *Programming Languages*

Expert systems can be written in conventional programming languages such as FORTRAN and PASCAL, and a few have been. However, the most popular and widely used programming languages for expert systems applications are LISP (LISt Processing) and PROLOG (PROgramming in LOGic). Symbol-manipulation languages like these are more suitable for work in AI since they lend themselves more readily to representing complex concepts and rule-based systems. A set of rules can specify how a program should react to changing data without requiring detailed advanced knowledge about the flow of control. In a conventional program, the flow of control and use of data are predetermined by the program's code.

(ii) *Shell Systems*

A shell is a package which can provide the means of building, using and maintaining a system without the need for developing it from scratch. Most shell systems are domain-independent (or empty systems). Their use provides great saving in development time but performance usually suffers. An example of shell systems is EMYCIN, or Empty MYCIN that has been generated from MYCIN. A broad and detailed survey of expert system building tools may be found in Harmon and King (1985).

2.4 Expert Systems Application Areas

The preceding sections gave an overview of expert systems structure, characteristics, expert system development procedure and of how they differ from that of conventional

programs, as well as discussion of the tools used to build an expert system. This section, first, examines the different categories of problems to which expert systems have been applied. It then presents an example of the early expert systems applications.

2.4.1 Application Areas

Expert Systems can be classified according to the type or category of problems they have been applied to, independently of the specific application field.

Typical applications and the corresponding categories of problems that have been found good applications for expert system technology are reported below (Hayes-Roth et al., 1983):

Predicting:	Inferring probable outcomes from given situation.
Monitoring:	Examining realtime data and watch for developing malfunction.
Designing:	Configuring objects on the basis of specifications and constraints.
Instructing:	Providing problem simulation and decision checking and help individualize instruction by diagnosing learner weaknesses and prescribing remedial lessons.
Diagnosing:	Infer malfunctions from a set of observable symptoms.
Interpreting:	Infer situation descriptions from sensor data.
Controlling:	Interpreting, predicting, repairing and monitoring behaviours.
Planning:	Evaluate possible future actions to determine the most logical series of steps leading to a desired goal.
Repairing:	Combine the diagnosis and planning systems to create and execute plans for repairing faulty systems.
Debugging:	Evaluate source code to identify syntax errors and predict errors in program logic based on defined goals.

2.4.2 Characteristics of tasks Suitable for Expert System

Not all fields of knowledge are currently suitable for expert system applications. It is therefore important to analyse the main attributes that make an application domain appropriate for an expert system approach. What should one look for in a task in terms of its suitability for expert system technology?

Expert systems are applicable to situations in which analysis, judgement, and experience are critical to an effective solution. There are many problems for which no efficient algorithm exists or for which the data required to solve the problem by an algorithmic method is missing or unobtainable. Expert systems often work well in these situations because they rely on rules of thumb, models, and general problem-solving techniques to provide problem solutions.

For a task to qualify for expert system technology, one or more of the following prerequisites must be met (Prerau, 1985; Waterman, 1986):

(a) First, there should be one or more articulate experts in the task domain that are able and willing to contribute their time and energies on behalf of the project.

(b) There are books or manuals or other written materials discussing the domain and the specific problem.

(c) The problem should be of narrow focus. That is to say, it should have a finite domain from which the knowledge base can be coded.

(d) The task domain should neither be too easy (not worth to be faced with a sophisticated computer-based system) nor too difficult (beyond the capabilities offered by current technology).

(e) Expertise is too limited, perishable, or unstructured to apply in a cost-effective manner.

(f) The problem is one for which no efficient algorithm exists. It requires symbol manipulation and heuristic solutions.

(g) There should be a minimum of common sense knowledge involved in the decision, and preferably none.

(h) The task should be primarily cognitive in nature; that is, successful performance is not based on physical ability or common sense.

(i) Finally, the completed system is expected to have a significant and measurable pay-off.

Sometimes an expert system can be built that does not exactly match these prerequisites; for example, the abilities of several experts rather than one, might be brought to bear on a problem (Fu et al., 1987).

2.4.3 An Example: Application to Medical Diagnosis

Expert systems technology has been applied to many different types of problems, as discussed above. Diagnosis has been one of the earliest applications areas as well as being one of the most important and interesting.

Numerous expert systems have been developed for diagnostic reasoning, many of them in the medical area. Most of the earlier successful systems were rule-based programs, like MYCIN (Shortliffe, 1976, Buchanan and Shortliffe, 1984), which is among the best known. For this reason some of its features are described below.

MYCIN Features

MYCIN is an expert system developed by Buchanan and Shortliffe in the early 1970's as a part of the Heuristic Programming Project at Stanford University. MYCIN is designed to help diagnose bacteriological blood infections. MYCIN is a rule-based expert system; that is, knowledge in MYCIN is represented as a series of production rules. The MYCIN system uses a backward chaining, or goal oriented control structure. The rules are stored internally in LISP. The following are excerpts (English version) from MYCIN used as demonstration (from Buchanan and Shortliffe, 1984).

Patient Information

MYCIN begins by requesting some basic information about the patient, as follows.

1) Patient's name: **PT538**

2) Age: **34 YEARS**

3) Sex: **MALE**

Diagnosis

Next, MYCIN begins to collect medical information by asking questions such as the following about test results, current treatment, and symptoms.

6) Please enter the results of pending cultures in the following table:

	culture	date	stain
site	number	collected	examined
CSF	**1234**	**30-JAN-77**	**YES**

10) Do you suspect PT538 may have an infection at a site from which you have not obtained culture specimens?

NO

11) Please enter information on current antimicrobials.

drug	route of	date

name	administration	started
CEPHALOTHIN	**IV**	**23-JAN-77**

15) Has PT538 recently had objective evidence of abnormal neurologic signs (nuchal rigidity, coma, seizures, etc.) documented by physician observation or examination?

YES

The CSF cultures will be considered to be associated with MENINGITIS.

After asking 15 questions (not all of which have been reproduced here), MYCIN has made a diagnosis of meningitis.

Causes

Next, MYCIN attempts to determine the causes of the infection that it has diagnosed.

17) Does PT538 have an abnormal chest X-ray?

NO

26) Has PT538 been exposed to any contagious disease recently (e.g. meningococcal disease, mumps)?

NO

27) Please enter CSF findings in the following table.

CSF WBC	Percent PMN	Protein	CSF Glucose	Peripheral Glucose
100	**99**	**280**	**5**	**UNKNOWN**

Notice that one of the answers to question 27 indicates that not all the information sought by MYCIN is available.

35) Is meningitis a hospital-acquired infection?

YES

39) Is PT538 a burn patient?

YES

I have found evidence (RULE545 RULE 557) that treatment should cover for E.coli causing meningitis.

I have found evidence (RULE 545 RULE 557 RULE 578) that treatment should cover for Pseudomonas-aeruginosa causing meningitis.

I have found evidence (RULE 545 RULE 557) that treatement should cover for Klebsiella-pneumonia causing meningitis.

name	administration	started
CEPHALOTHIN	**IV**	**23-JAN-77**

15) Has PT538 recently had objective evidence of abnormal neurologic signs (nuchal rigidity, coma, seizures, etc.) documented by physician observation or examination?

YES

The CSF cultures will be considered to be associated with MENINGITIS.

After asking 15 questions (not all of which have been reproduced here), MYCIN has made a diagnosis of meningitis.

Causes

Next, MYCIN attempts to determine the causes of the infection that it has diagnosed.

17) Does PT538 have an abnormal chest X-ray?

NO

26) Has PT538 been exposed to any contagious disease recently (e.g. meningococcal disease, mumps)?

NO

27) Please enter CSF findings in the following table.

CSF WBC	Percent PMN	Protein	CSF Glucose	Peripheral Glucose
100	**99**	**280**	**5**	**UNKNOWN**

Notice that one of the answers to question 27 indicates that not all the information sought by MYCIN is available.

35) Is meningitis a hospital-acquired infection?

YES

39) Is PT538 a burn patient?

YES

I have found evidence (RULE545 RULE 557) that treatment should cover for E.coli causing meningitis.

I have found evidence (RULE 545 RULE 557 RULE 578) that treatment should cover for Pseudomonas-aeruginosa causing meningitis.

I have found evidence (RULE 545 RULE 557) that treatement should cover for Klebsiella-pneumonia causing meningitis.

I have found evidence (RULE 507) that treatment should cover for Diploccus-pneumonia causing meningitis.

Treatment

My therapy recommendations will be designed to treat for organisms that either are very likely or, although less likely, would have a significant effect on therapy selection if they were present. It is important to cover for the following possible infection(s) and associated organism(s):

INFECTION-1 IS MENINGITIS

+ <ITEM-1> E.COLI

+ <ITEM-2> PSEUDOMONAS-AERUGINOSA

+ <ITEM-3> KLEBSIELLA-PNEUMONIAE

+ <ITEM-4> DIPLOCCUS-PNEUMONIAE

41) Does PT538 have a clinically significant allergic reaction to any antimicrobial agent?

NO

43) PT538's weight in kilograms:

7 0

RECOMMENDATION-1

My preferred therapy is as follows:

In order to cover for items <1 2 3 4 >:

give the following in combination:

1) AMPICILLIN

Dose: 3.5g q4h IV

2) GENTAMICIN

Dose: 119 mg q8h IV plus consider giving 5mg q24h Intrathecal.

Comment: Monitor serum concentrations.

Since high concentrations of penicillins can inactivate aminoglycosides, do not mix two antibiotics in the same IV bottle.

Explanation Facility

MYCIN can explain *why* it asked a particular question and *how* it reached a particular conclusion.

For example, if the user asks MYCIN to explain the relevance of the information in question 35, the explanation facility presents its reasoning as follows (probabilities of correctness or "certainty factors" are attached to each conclusion):

35) Is meningitis a hospital-acquired infection?

WHY

(2.0) This will aid in determining the organisms (other than those seen on cultures or smears) which might be causing the infection. It has already been established that (2.0) the infection which requires therapy is meningitis, and

(2.2) organisms were not seen on the stain of this pending CSF culture, and

(2.3) the type of the infection is bacterial, and

(2.4) neurosurgery is not one of the relevant items from the history of PT538, and

(2.5) the age of PT538 is greater than 10 days

Therefore, if

(2.6) the infection was acquired while the patient was hospitalized

then

there is evidence that the organisms (other than those seen on cultures or smears) which might be causing the infection is e.coli (.75) staphyloccus-coag-pos (.3) pseudomonas-aeruginosa (.3) Klebsiella-pneumonia (.5).

When MYCIN is asked "how" it reached a conclusion, it lists the production rules that it used and the certainty factors that emerged from the successive use of the rules.

2.5 Expert Systems Applications to Safety and Reliability

2.5.1 Possible Expert Systems Application Areas

Application of expert systems to safety and reliability improvement is a new area. Use of expert systems for nuclear process plants safety/reliability covers a large range of areas, as described in the next section. However, their use for major hazard process plants safety/reliability has been restricted to few areas. Andow and Fergusson in a recent paper on *"Applications of Knowledge-based Systems in Chemical Process Safety"* (Andow and Fergusson, 1987), state that expert systems are likely to be applied to a wide range of safety problems in process plants. They identified several possible ergonomics/human factors and safety-related applications of expert systems techniques in chemical process plant design and operation as illustrated by Table 2.1. The authors concluded that:-

Table 2.1 Possible Applications of Expert Systems in Process Safety

	A. Applications in Plant Design
(a)	Protective system specification
(b)	Hazard and operability studies (HAZOP)
(c)	Equipment selection
(d)	Control system design
	B. Applications in Plant Operation
(a)	Fault diagnosis
(b)	Complex operating sequences (of valves and pumps)
(c)	Intelligent control systems.

(i) Some of these applications, such as equipment selection, were considered possible, but difficulties in implementations should not be underestimated. These include: knowledge elicitation, system size, system design methodology, and software tools.

(ii) Some applications, such as HAZOP, appeared to be too ambitious at that time (in 1987).

2.5.2 Prior and Current Work

During the last 4-5 years, rapid developments have occurred in the development of AI and expert systems techniques to aid operators, engineers and managers in decision

making related to risk and safety, particularly in nuclear industries where various systems have been or are under development.

Table 2.2 shows a listing of the AI applications described at the *"First International American Nuclear Society Topical Meeting on Computer Applications for Nuclear Power Plant Operation and Control"*, held in September 1985 in the USA.

Some of the attempts that have, or are being made, to apply expert systems technology to safety and reliability problems in nuclear and process industries are briefly outlined below. Further details can be found in the references cited.

A. HAZARD IDENTIFICATION AND ASSESSMENT

1. **HAZTCHECK** - Check-list Approach for Hazard Identification

HAZTCHECK is an expert system being developed, as its authors (Reeves et al., 1988) claim, in order to progress the check-list approach to hazard identification. This program is written in PROLOG and runs on IBM-PCs and compatibles. The program permits the ready probing of a check-list and its update.

2. Pressure Relief Valve Selection

This project is aimed at developing an expert system for the specification and selection of pressure relief valves (Miller, 1988). It is part of the work carried out by the Expert System Group at BHRA to build large consultative expert systems. One of the strategies for the use of the expert system will be its contribution to plant hazard analysis.

Table 2.2 AI Applications for Nuclear Power Plant (NPP) Operation and Control

Application	Author(s)	Place
1. A knowledge based system for plant diagnosis	Kigushi et al.	Japan
2. An event-driven control method for automatic operation and control of NPPs	Kinoshita et al.	Japan
3. Semantic network approach to automated failure diagnosis in NPP.	Washio et al.	Japan
4. Development and verification of an accident diagnostic system for NPP by using simulator	Yoshida	Japan
5. Expert system for real-time diagnostics and control	Deegan et al.	USA
6. Application of AI to improve plant availability	Frank et al.	USA
7. Expert systems in real-time environments	Moore et al.	USA
8. An AI program in a computer application supporting nuclear reactor operations.	Stratton et al.	USA
9. Control system verifier using automated reasoning software.	Smith et al.	USA
10. An AI approach to sensor conflict detection.	Chandrasekan et al.	USA
11. Causal representation and explanation of NPP operation.	Underwood et al.	USA
12. A reactor safety assessment system (RSAS).	Sebo et al.	USA
13. The feasibility of using simple intelligent system as aids to plant operators.	Abbot et al.	Belgium
14. AI methodologies application to construction management decision support systems.	Walmsley	USA
15. A prototype fuel-shuffling system using a knowledge-based tool kit.	Fraught et al.	USA

3. **HAZOPEX** - Hazard and Operability Study

HAZOPEX is an expert system being developed at the Occupational Safety Engineering and Electrical Engineering Laboratories of the Technical Research centre of Finland. HAZOPEC aims to give support to process designers in the evaluation of new process systems, to reduce the resources and time needed for HAZOP analyses, and to assure the quality of HAZOP analyses (Suokas et al., 1987). HAZOPEC includes the necessary methodological knowledge and the user is required to have a good knowledge of the process system to be investigated.

4. **HASTE** - Hazard Assessment

HASTE (Hazard Assessment System for Toxic Emissions), developed by Environmental Research and Technology (Concord, Mass.) (Hushon, 1986), is one among other expert systems that have been developed for use at plant site to determine the location of an emergency and to predict danger to the environment and community health that could result from atmospheric chemical releases.

B. SYSTEM SAFETY AND RELIABILITY

1. **ESYRE** - Expert System in Reliability

ESYRE is an expert system in reliability implemented on a personal computer. Its developers (Popchev and Zlatareva, 1985) claim that it is designed for the broadest spectre of users (scientists, engineers, constructors, specialists in the field of standardization of reliability, students, etc.). Its problem domain has been defined as consisting of:

(i) general problems of reliability;

(ii) choice of criteria and standardizing the level of reliability of products;

(ii) computation, evaluation, optimisation, forecasting and modelling of reliability.

2. **EXPRESS** - System Safety Studies

EXPRESS is an expert system developed by Electricite de France (EDF), a French utility, for the automation of reliability studies (Ancelin, 1987). EXPRESS is designed to help the analyst to build fault trees for the study of safety systems reliability. The knowledge representation of the system is based on two types of inference engines ALOUETTE (EDF-DER, 1984), and LRC (EDF-DER, 1985). The system was first applied to the construction of fault trees for static thermohydraulic systems in the Paluel nuclear power plant. The second phase of the project is being implemented (Ancelin et al., 1987) and involves building a fault tree for electric power systems.

3. **ERNEST** - Fault Tree Construction

The construction of fault trees for reliability computation is time-consuming and affected by several sources of potential error. To ease part of these difficulties an expert system (ERNEST) has been designed. ERNEST is an expert system for the interactive knowledge-driven construction of fault trees (Garriba et al., 1987). The system is at present under testing with the help of tutorial case-studies.

4. **SQUIMP** - Event Tree Sequence Calculation

SQUIMP has been developed for event tree sequence importance calculations (Dixon et al., 1985). SQUIMP provides for prompted data entry, generic expansion and on-line pruning of the event trees, Boolean reductions, and importance factor selection. It is intended for applications in the context of probabilistic risk assessment (PRA).

5. **PC-Predictor** - MTBF and FMECA Analysis

5. **PC-Predictor** - MTBF and FMECA Analysis

PC-Predictor is an expert system designed to analyze MTBF (Mean Time Between Failure) and FMEKA (Failure Mode Effects Criticality Analysis) of computer system (Zemva et al., 1987). PC-Predictor is a basic version of a Mainframe program which permits the evaluation of system reliability according to stress analysis method of MIL-HDBK-217/Notice1. Predictor performs the reliability analysis and FMECA simultaneously. A perquisite to FMECA is reliability data.

6. Expert Interface to Reliability Data Bank

The effective use of technical data banks is difficult and requires a skilled specific experience. A feasibility study and a preliminary design have been carried out (Carlesso et al., 1987), concerning the development of an expert interface to ERDS (European Reliability Data System). The implementation of the system is in progress. The project has been divided into two phases: design and implementation of prototypes, and design and construction of the target system.

C. EXPERT SYSTEMS for HUMAN RELIABILITY IMPROVEMENT

Information management and display of relevant data concerning plant operation has been a concern of the nuclear and process industries from their beginnings. Since the incident at TMI, this matter has had much careful scrutinity. Hayes-Roth's tutorial (Hays-Roth, 1984), on knowledge-based expert systems states that computer-stored skills of specialists may be used to aid operators in solving problems. A major benefit from such systems is the consistent applications of logic and timely solutions to problems.

The nature of the operator's functions in process plants is changing from manual operations to control and decision making, such as identifying, comparing, analysing/ interpreting, evaluating/ selecting, planning, and verifying. These cognitive activities or *"intensive functions"*, require knowledge processing of the operator. Whereas the reliability of these actions may deteriorate if the conditions are unfavorable, they cannot easily be automated. But expert systems techniques may assist the operators in these tasks. Several efforts are being made to develop expert systems that aim to improve human reliability by supporting operators decision making. Following is a description of some particular project that are relevant for the application of expert systems in nuclear and process industries

1. Expert Display System

Beltracchi (1988) from the U.S. Nuclear Regulatory Commission describes an expert display system that controls automatically the display of segments (a combined series of graphic elements) on a cathode ray tube's screen to form an image of plant operations. The image consists of an icon of: 1) the process, 2) plant control systems, and 3) safety systems. A set of data-driven, forward-chaining computer stored rules control the display of segments. As plant operation changes, measured plant data are processed through the rules, and the results control the deletion and addition of segments to the display format. The icon contains information needed by control rooms operators to monitor plant operations.

2. **CEALMONT** - Emergency Action Level Monitor (EAL)

CEALMONT is a rule-based emergency action level (EAL) monitor implemented on an IBM-PC (Touchton et al., 1985). It helps to automate EAL classification procedures in a real-time processing environment. The knowledge base includes the logic of current

EAL tables and higher-level rules to resolve ambiguities and data conflicts, identify false alarms, and draw inferences in the event of missing data.

3. **RSAS** - Reactor Safety Assessment System

An expert system called RSAS has been developed for the U.S. Nuclear Regulatory Commission (NRC), to assess the status of a reactor system during an accident and recommend corrective actions to reactor operators (Sebo et al., 1985). RSAS performs data consolidation and consistency checks, monitors safety setpoints, and determines diagnostic hypotheses related to core damage and containment releases.

4. **VIOLET** - Rapid Fault Diagnosis

In hazardous situations, the most important human operators'task and responsibility is to avoid major faults and to diagnose any faults that occur as rapidly as possible. A major problem in that operator's task is the interpretation of the information which is measured. Expert systems techniques can be used to help operators in condition monitoring of process and rotating machinery. By forecasting the changes in the plant, many faults can be predicted and avoided. One example of expert systems developed for that purpose is VIOLET (Milne, 1988). VIOLET is an on-line expert system applied to non-destructive testing for interpreting vibration data.

5. Safety Integrity Management

The system proposed by Andow (1988), is intended to improve safety by enhancing the reliability of operator responses as they are faced with ever more complex systems. The system would be able to monitor protective systems and warn the operator of safety system degradation.

5. GRADIENT - Control Room Operator Decision Aids

GRADIENT is a five year project, in the European ESPRIT programme, to design and develop a graphics and knowledge based dialogue interface for industrial supervision and control (S&C) systems (Hollnagel, 1987).

The aim of the GRADIENT project is to investigate the use of knowledge based systems to support operator decision making during normal and adverse conditions of system operation. This support will be supplied by a set of expert systems which will provide:

- Real-time and fault identification and alarm handling;
- Prevention of incidents through monitoring of operator actions;
- Increased flexibility of the graphics interface through abandoning the restrictions of pre-defined displays.

The GRADIENT system contains the following subsystems:

- QRES.- Quick response system. Supports the operator during a system failure.
- RESQ.- Monitors and evaluates the operator's actions.
- SES - Supports the operator with knowledge and information.
- DIS - Coordinates dialogue between the operator and both the Process and SES.
- GES - Graphical system. It will dynamically compose pictures and pictures sequences, using knowledge of the process, the user model, graphical representation techniques and dialogue techniques.

By making the required operator actions less knowledge intensive, the reliability of the operator's performance is expected to increase.

2.6 Some Expert Systems Limitations

Despite all their capabilities, expert systems do have limitations. Some of the limitations to expert system development may originate from technology inheritance, environment, and cost (David Hu, 1988).

Because expert system technology is still evolving, limitations include inherent shortcomings such as narrowness of expertise, inability to recognize knowledge boundaries, limited explanation facilities, and difficulty in validation.

Because building and maintaining a large knowledge base requires substantial effort, most expert systems cover a narrow range of expertise. Part of the reason is due to current computing facilities which limit the speed and capability of search in expert systems. Even when an expert system achieves a broad coverage of knowledge, it becomes shallow in representing associations between elements in the knowledge base.

One major limitation is concerned with the inability of expert systems to recognize their knowledge boundaries. Most expert systems do not deal competently with problems at the boundaries of their knowledge. They do not have the knowledge built in to determine when a problem is beyond their capabilities or outside their fields.

Expert systems are limited by the information in their knowledge base and by the incremental and iterative process by which knowledge is elicited and entered. They cannot report conclusions that are not already implicit in the knowledge base. The process by which knowledge is elicited, entered and tested is likely to produce inconsistencies and incomplete knowledge bases; hence an expert system may unexpectedly exhibit knowledge gapses. Expert systems may need constant

maintenance to reduce mistakes in complex or derived cases that are not fully represented in their knowledge base.

Expert systems cannot yet replace a human expert completely because the expert is not someone who is just following the rules but who has the experience to know an exceptional case and what to do about it. In other words, the expert knows when to break the rules.

The explanations are often primitive and a human expert may need to explain again what the expert system has explained.

Although validation of software programs is time consuming, the effort required to validate expert systems is many factors greater than that for conventional software programs. Communication between human experts and knowledge engineers in identifying and correcting mistakes in an expert system can be a formidable task.

The environment in which an expert system is developed is significant. Two of the potential limitations that exist within an environment are hardware and software. Computer hardware may be slow, not equipped for symbolic processing, or expensive. Except for large, expensive software tools, expert system tools have limited knowledge representation methods.

Cost is a major source of problems for developing expert systems. To build an expert system, the knowledge engineer extracts the requisite knowledge from human experts or other sources and laboriously implements it into the knowledge base. This effort is time consuming and, at present, the knowledge engineers are in short supply.

2.7 CONCLUSION

This chapter has provided an overview of expert systems technology and applications. First, it has explained what expert systems are, how they work, how they differ from conventional techniques, how, and with what tools, to build them, and what types of problems they have been applied to. An example of one early developed rule-based expert system called MYCIN was provided to illustrate application of expert systems technology to medical diagnosis. Finally, examination of some expert systems being developed and used in risk and safety management of nuclear and process plants was presented.

The main conclusions are that expert systems provide a means of making the knowledge and expertise of skilled staff available to others. Many standalone consultative expert systems have already been implemented using microcomputers. Expert systems applications include medical and equipment diagnostics, design assistance, equipment configuration, system safety/reliability analysis.

As with human experts, these systems are limited by their knowledge and 'understanding' (i.e.model) of the application. The problem of eliciting this knowledge from humans and representing it in a structured validatable form is a major barrier to the wider use of expert systems, e.g., in large or safety-critical applications, and is the subject of much research and development activity. In the meantime, expert systems can be useful, with human supervision, as intelligent assistants.

The promise of expert systems is immense, but a word of caution is in order, too. Expert systems have been around for a long time and have had some significant impact, but they are not by any means a mature technology. Yet, in spite of that warning, the hopes are high. As the technology matures and costs decline, the use of expert systems can be expected to spread widely. Over the long term, expert systems

preserving know-how, distributing knowledge more effectively, and freeing up experts from time-consuming tasks.

Expert systems for risk and safety management have been developed and are currently being used. They have proved to be a very useful tool for solving many safety-related decision-making problems in nuclear process plants including failures and accidents diagnosis, fault and event trees constructions, and system reliability assessment.

Expert systems have a wide range of potential applications to chemical process plants design and operation, including hazard identification, equipment selection, fault diagnosis and process control.

Efforts have already begun on a number of these applications. They are being used, for example, to assist in hazards identification and assessment at chemical plants to support process designers in the evaluation of new process systems, and predict the dispersion of hazardous materials in the event of a release.

The results of the literature review of expert systems applications, discussed in this chapter, show clearly that their use for safety improvement is increasing. Human reliability is an important determinant of safety and reliability improvement. However, the use of expert systems for safety in nuclear and process plants was mainly restricted to hazard identification, hazard consequences assessment, system reliability analysis, and decision aids to improve human reliability. But nothing had, as far as the author was aware, been done on human error/reliability modelling and quantitative prediction.

The main characteristics of the human reliability analysis problem domain as part of probabilistic safety/risk assessment studies are discussed in the next chapters.

CHAPTER 3

SAFETY IN NUCLEAR AND MAJOR HAZARD INDUSTRIES
Risk Quantification and Human Reliability Analysis

3.1 Introduction

Nuclear, chemical and process industries have been subject to a number of serious accidents involving major fires, explosions, and toxic and radioactive releases. Some of these accidents have led to heavy loss of life and did extensive damage to property. Some significant industrial accidents are listed in table 3.1.

Because of the large quantities of hazardous substances they handle, and the potential they have to put employees and the general public at significant risk, chemical and process industries have been classified as major installations.

Nuclear, chemical and process industries operate plants that have the potential for large accidents which could result into major loss of life and economic consequences. There is also an increasing public concern about the safety of these large-scale installations. These industries share, therefore, a common imperative to prevent major accidents which could harm the public, produce major financial losses and result in an adverse public opinion.

In the early 1970s, the process industries began to realize that with new higher intensity, higher inventory processes, the practice of learning by mistakes in the field of safety was no longer tenable. Much work was carried out to develop methods for the identification of what could go wrong and for assessing the likelihood of such undesired events. These techniques were developed primarily as aids to decision taking, to help managers ensure that appropriate resources were applied in the areas

Table 3.1 Some significant accidents in nuclear and process industries (Thompson, 1987).

Place	Year	Event	Consequences
Brockton, Massachusetts	1905	Boiler Explosion	58 deaths
Oppau, Germany	1921	Ammonium nitrate explosion	430 deaths
Zarnesti, Rumania	1939	Toxic release (chlorine)	60 deaths
Cleveland, Ohio	1944	LNG pool fire	130 deaths
London, UK	1952	Smog (atmospheric pollution)	c. 4000 deaths
Uskmouth, UK	1956	Steam turbine explosion	2 deaths
Windscale, UK	1957	Nuclear reactor fire	c. 17 deaths* (central estimate)
Potchefstroom, South Africa	1973	Ammonia storage tank failure	18 deaths
Flixborough, UK	1974	Vapour cloud explosion	29 deaths
Scunthorpe, UK	1975	Foundry steam explosion	11 deaths
Seveso, Italy	1976	Toxic release (dioxin)	No deaths recorded
Los Alfaques, Spain	1978	Propene flash fire	215 deaths
Three Mile Island, Pennsylvania	1979	Nuclear reactor LOCA	c. 1 delayed death* (central estimate)
Ixhuatepec, Mexico	1984	LPG explosions	c. 500 deaths
Bhopal, India	1984	Toxic release (MIC)	c. 2500 deaths
Chernobyl, USSR	1986	Nuclear reactor fire	31 deaths

* The estimates for the developments of cancers due to these accidents are determined on the conservative assumption that there is a linear dose - risk relationship, with zero additional risk at zero dose.

were they would do the most good.

By the mid-1970's, there was already concern over the inadequacy of the then current mechanism for the control of industrial activities which had the potential for causing incidents which might have a major impact on the health, safety and property of the general public. During this period the method of ensuring safety was usually an empirical approach, based on previous experience and engineering judgement. Nowadays, this is not enough. Rasmussen has pointed out (Rasmussen, 1988) that one important implication of,

> *" The trend towards large scale industrial process plants and the related defence-in-depth design practice"*

is that,

> *"...the actual level of safety cannot be directly controlled from empirical evidence. For hazardous large scale installations, design cannot be based on experience gained from accidents, as it has been the case for accidents in minor separate systems....The days of extensive pilot plant tests for demonstration of the feasibility of a design is over and safety target has to be assessed by analytical means based on empirical data from incidents and near misses, i.e., data on individual, simple faults and errors. Consequently, for industrial process plant, large efforts have been spent on developing methods for Probabilistic Risk Analysis."*

Another factor, in addition to the trend towards larger plants and defence-in-depth, is the increased public awareness about the safety of the process industries especially after the recent major accidents such as Flixborough 1974, Three Mile Island 1979, Bhopal 1984 and Chernobyl 1986. Human operators have been found to have played a major part in the occurrence of these accidents. It is, therefore, becoming a necessity for effective management and control of risk to involve a systematic analysis of human reliability. The HSE has recently (HMSO, 1989) published a guidance booklet on

"Human factors in industrial safety" which urges managers to consider human factors as a distinct element of their everyday work which they must recognize, assess and control if they are to minimize risk.

These changes have resulted in the emergence of new risk criteria and regulations concerning the safety of hazardous installations (for example, the EEC's Seveso Directive, 1982, the UK's CIMAH Regulations and the HSE's document on "The tolerability of risk from nuclear power stations", 1987). The regulations now require the owners of particularly hazardous plants to write, and present to the regulatory authorities, a safety case report to demonstrate that they are aware of hazards and they are taking adequate precautions.

The regulatory authorities are faced with many types of decisions in discharging their legal responsibilities for the realization of nuclear and process plants. These may be categorized as follows:

1. How safe should plants be?
2. How safe are they?
3. Does the safety of plants need to be improved?
4. How should the desired level of safety be ensured during the lifetime of the plant?
5. What issues require research to improve the state of knowledge and enhance effective regulation?

The important role that probabilistic or quantified risk assessment (PRA) plays as input to policy decisions about the potential risks from nuclear and major hazards, has been described in a paper recently published by the HSE, *"Quantified risk assessment: Its input to decision making"*, (HSE, 1989a).

The HSE's paper emphasizes that PRA is an indispensable element in predicting the many possible hazards that can arise in a complex plant, nuclear or non-nuclear, and in assessing how likely or unlikely each eventuality is, so as to see how best to control and if possible reduce the risks that are identified.

This chapter briefly examines how historically the safety questions have been answered in process control - particularly nuclear - industries (for more detail see Tanguy, 1988). It then discusses the purpose and content of the techniques of quantified or probabilistic risk assessment (QRA or PRA). From the discussion of the techniques a number of conclusions will be drawn about the applications and usefulness of PRA as a tool for safety and reliability improvement, as well as some of its limitations. Finally, inclusion of human reliability analysis (HRA) within the overall framework of PRA is discussed. For a more detailed description and discussion of the techniques of PRA, see the *"PRA Procedure Guide..."*, (NUREG/CR-2300, 1983) and the*"Probabilistic Risk Assessment (PRA) -- Status Report and Guidance for Regulatory Application."*, (NUREG-1050, 1984).

3.2 Safety Development in Nuclear Industries

Nuclear power has developed very fast since the first commercial stations were built in the early fifties. However, throughout the world, opposition to nuclear power is growing. Since the accident at Three Mile Island in 1979, no electrical utility has ordered a nuclear reactor in the United States (Klueh, 1986). Despite its significant advantages: cheaper electricity from nuclear energy and less threat to the environment than from other sources and good safety record, public opinion shows great concern over nuclear power and the safety of nuclear power plants has become the focus of the international nuclear community since the Chernobyl accident in 1986.

Nuclear power plants accidents differ from those in nonnuclear power plants because they can potentially release significant amounts of radioactivity to the environment. By far the largest amount of radioactivity resides in the reactor core. The fuel is subjected to heating due to absorption of energy from the radioactive decay of fission products.

The nature, throughput and inventory of radioactive materials involved in these process plants necessitates very high standards of safety in their design, construction and operation.

Although, as Hinton (1957) pointed out, *"all other engineering technologies have advanced not on the basis of their success but on the basis of their failures"*, the nuclear industry could not afford to do likewise, because of its associated hazards. (Indeed, as the scale of other technologies has increased, many technologies now have the potential to cause unacceptable damage, and 'progress through failure' is seldom nowadays justifiable.) Nevertheless, the nuclear industry has learned much from a number of non-catastrophic accidents, notably the Windscale fire in 1957 and the TMI accident in 1979. Much more will be learned from the Chernobyl accident in 1986.

3.2.1 Early Development

Safety has been an important consideration from the early development of nuclear reactors. Containment for protection of the general public was, and is still today, one of the central issues in reactor safety assessment. The first proposal for a contained reactor was put forward to "The Reactor Safeguards Committee" in 1947 (Okrent, 1947).

The nuclear industry grew rapidly in the years following the Second World War. The Windscale accident in 1957 in the UK was the first, and until Chernobyl, the only one

of its time, large-scale radioactive release into the environment with its potential for long-term consequences.

The 1957 WASH-740 accident analysis (revised many years later in 1966) was the first report which gave a quantitative estimate of the maximum conceivable consequences of a severe uncontained reactor accident. It became the basis for the liability limits to be included in the Price Anderson Act, which defines the provisions of insurance of nuclear power plants in the US. WASH-740 represented the main reference on what could be the consequences of a very severe nuclear accident until the Rasmussen report (WASH-1400, 1975), and until the Chernobyl accident in 1986.

3.2.2 Safety of design: 1957-1967

The dominant safety aspect of this period is the importance given to safety of the design. Most of the concepts which are still in use were established by that time including the main safety functions: controlling the chain reaction; cooling the core; and containing the radioactive materials. The concept of defence-in depth, with the requirement of redundancy to fulfil the single failure criterion, and of postulated initiating events to give the design basis for the safety features, were established. The question of pressure vessel integrity was raised in reference to the safety approach, based on prevention of accidents and on mitigation of their consequences should they occur.

In parallel, significant improvements were brought to codes and standards, from stress analysis to in-service inspection.

Reactivity transients were the subject of considerable investigations, in particular as regards fuel failure mechanisms, in many research facilities.

Safety by siting was introduced during this period - it is safe if it is sited - and difficulties in the use of precise power reactor site criteria, such as population and dose limits, became more and more apparent.

In the American Energy Committee (AEC) approach, the concept of "maximum credible accident" was used, presented for the first time in 1959. Although it played a major part in safety assessment, it was not universally accepted (Tanguy, 1988). It is defined as:

> *"the upper limit of hazard, i.e. fission product release, against which features of the site must be compared"* (Beck, 1963).

There was, at this time, a slow move towards the development of realistic safety criteria as Farmer noted in 1964:

> *"It was recognized that the limitations placed on a reactor at the design stage.... should be related to real and recognizable criteria which could be tested"*

And there was a move away from the simple concept of *"safe"* and *"unsafe"* to the recognition of risk or *"how safe is safe enough"*.

3.2.3 Safety of Construction: 1967-79

During the second period, from 1967 to TMI, the emphasis was on safety of construction. One key safety aspect was introduced at this time: Quality Assurance. The importance of safety during the construction stage has always been recognized, however, very little guidance was available. A lot of guidance was given later, and quality assurance was its main source.

Apart from quality assurance, safety design underwent a considerable evolution during these years. In 1970, a programme of safety guides (later renamed regulatory guide) was initiated by the AEC to implement the design criteria.

Finally, this period is also characterized by the publication of the Rasmussen report, WASH-1400, in 1975 which was followed by a general consensus on the benefits for safety which could be gained from a probabilistic approach, as a supplement to the deterministic one used in design.

On 28 March 1979, the safety scene looked satisfactory on the whole. The safety approach was coherent, and there were no pending serious issues. In fact, some members of the nuclear community were even convinced that NPPs might well be not only safe, but too safe.

Maybe at that time it was overlooked that NPPs had evolved over years, and had increased in power capacity. The decay heat levels were much higher. Engineered safety features were added to reduce the likelihood of accidents, but the designs had become more complicated. There were now important relationships between the possible failures of various safety features. And more important, most discussions dealt with design, while not enough attention was given to safety in operation and its human component.

3.2.4 Safety in operation: 1979-86

The third period includes the lessons learned from TMI. They were re-emphasized after Chernobyl. Only after TMI, operational safety and human factors were given the attention they deserved. Many essential safety aspects played a role in the TMI accident including:

1. the importance of adequate operating procedures;
2. the need for appropriate training for operating personnel;
3. the necessary improvement of the man machine interface;
4. the usefulness of operating experience feedback;
5. the requirement for efficient emergency plans; and
6. the danger of improper "mind-sets" at all levels of the operating organization.

Also in this period, probabilistic methodologies were at last used in practice to improve safety. "Safety goals" have to be used, even if implicitly, in the decision-making process. There was a conscious policy of trying to make nuclear power reactors safer than other industrial or technological entreprises. This general objective has been translated in terms of limited probabilities for harmful accidental consequences.

Finally, there is no doubt that this period has seen significant improvements in the safety of NPPs. The Chernobyl disaster does not necessarily contradict this statement, but it compels managers of these plants to proceed to a new and complete review of their safety philosophy and practices.

3.2.5 1986 and Beyond

No one can know what will be the main safety trends in the next decade in the technical field. The emphasis today in nuclear safety efforts is switching from establishment of standards and quality assurance to accident prevention through improved operational safety and accident mitigation, i.e., accident management, containment integrity, and emergency preparedness.

3.3 Development of Major Hazard Installations

3.3.1 Early Development

The process industry developed in connection with the industrial revolution, when human muscle power was substituted for machines. The first phase was ushered in at the end of the 18th century and early 19th century by the rapid growth of the textile industry which used inorganic chemicals such as chlorine. At the same time there was a growth in the manufacture of coal gas and coke making industry. Coal gas and coke oven gas and other nineteenth century fuel gases were highly toxic.

The chemical industry was ranked as one of the most unhealthy of the nineteenth century industries. During most of the nineteenth century the chemical industry was not exactly safety conscious [Kletz, 1988]. The period produced many human casualties. Air , water and land was polluted and the health of work people assailed.

Among the first industrial artifacts to give serious concern about public safety were boiler pressure vessels used to provide power and heat. Boiler explosions became an every day occurrence throughout industry as the underlying causes and mechanisms of failure were poorly understood and the boilers, of the fire tube type, often lacked safety valves and low level alarms.

3.3.2 Between the Wars Period

The chemical industry expanded greatly following the First World War. The early processes were batch operation. The advantages of continuous operation soon became apparent. The scale of chemical plants was increasing but methods of design and operation had not changed. The result was a series of fires and explosions, such as the Oppau disaster, Ludwigshafen, explosion of 21st September 1921 in which a store of some 4,000 tonnes of an ammonium nitrate mixture exploded, killing over 500

peoples. This accident may be said to have marked the beginning of the era of major chemical hazards (Marshall, 1987).

3.3.3 Post Second War period

The Second World War stimulated a significant acceleration in the rate of technological development, the results of which were applied for industrial and commercial purposes after the war. As a result, there have emerged new materials, new processes, even new whole industries such as the petrochemicals industry and especially of the techniques of handling liquefied gases. The handling of liquid hydrocarbons has been involved in some 40% of all serious incidents during that period (Marshall, 1987).

The post Second World War era has been strongly influenced by two factors. The first is the very strong commercial pressure to improve technological efficiency. The second is the growing public awareness of the high potential for disaster and effects on the environment of the new technologies which, on the other hand, have led to considerable improvement in their standards of living. The debate was particularly lively and is still about what level of risk is acceptable or unacceptable

The pressure for greater efficiency has led to increases in size of many traditional plants, typically by a factor of about ten, the grouping of associated processes in one area, and the building of new plants, larger than those built before, employing increasingly complex technology, and containing higher inventories of hazardous materials under more extreme conditions. This has brought about a very significant increase in the number of people who could be endangered at any one time should anything go wrong.

Complexity of process plants provides opportunities for errors in design, construction and operation. The pace of change in process industries associated with modern

technology allows less opportunity for learning by trial and error. It is increasingly necessary to seek to get design and operating procedures right first time. Because of their present-day size and throughput there are now many plants where a critical first mistake can result in disaster.

As Lees (1980) pointed out:

> *"The operation of such plants is relatively difficult. Whereas in the past chemical plants were small and could be started up and shut down with comparative ease, the start up and shut down of a large ...plant in an integrated site is a much more complex and expensive matter."*

The reliability of large plants has often been unsatisfactory. Compact layout has made maintenance more difficult. Human errors were very frequent particularly during emergency situations. There have been many examples of faulty equipment. Complete shut-down of the plant due to failure of a simple piece of equipment has been a frequent experience.

These factors have resulted in an increased potential for loss - both human and economic and delays in the commissioning and operation of large process plants.

Managers and designers realized the need for a new approach to safety of process plants, and *loss prevention*, which has its origin in insurance, was adopted as an approach to safety assessment in process industries. Loss prevention is concerned, among others, with identification of new hazards that arise out of new technology, hazards that can, or could cause, damage to plant and loss profit as well as those that cause, or could cause, injury. The main analytical technique used to identify hazards and review design in process industries is HAZOP (Hazard and Operability Studies).

3.3.4 Control and Automation of Process Plants

Modern processing plants store large inventories of hazardous materials that have to be kept under control. These plants depend greatly on sophisticated automatic equipments for their control and protection.

The processes were originally controlled and supervised manually. Some control and supervising functions were gradually automated when equipment for measuring flows and quality were developed. Application of electronic instruments for process control started in the beginning of the fifties. Implementation of control systems in terms of remote sensors and actuators and a central control room with PID (piping and instrumentation diagram) regulators became the general solution to process automation.

The thought of using digital computers for process control emerged in the mid-fifties. The computers used during the early period were large, slow, expensive, and unreliable. Because the computers were so unreliable, they had to be used in *supervisory modes* only, while ordinary analog controllers were used for the primary control functions. Two different approaches emerged. In the *operator guide* mode, the computer simply gave instructions to the operator about set points for the analog controllers. In the *set point control* mode, the set points were adjusted automatically from the computer. The major tasks of the computer were to find good operating conditions, to perform scheduling and production planning, and to give reports about production, energy, and raw-material consumption. It became clear that the demand for fast response to external events imposed special requirements on computer architecture and software. It was also found that many sensors were missing.

In 1961 Imperial Chemical Industries (ICI) programmed a Ferranty computer to do, in sequence, the control calculations normally done by conventional controllers. The

computer controlled the process directly. An operator communication panel can replace a large wall of analog instruments.

Advances in integrated circuit technology led to the development of cheaper, smaller, faster, and more reliable *minicomputers*. The development of minicomputers gave rise to a rapid increase in applications of computer control.

The development of the microcomputer in 1972 has had far-reaching consequences. With the microprocessors, it was possible to switch technology for the functions that were originally made by relay systems. Special-purpose programmable logic computers (PLC) appeared for realization of the logic and sequencing functions. The operator communication has been vastly improved in these systems by using colour video graphic displays. The first distributed computer-controlled system was announced by Honeywell in 1975. Hierarchical control systems with a large number of micros have been constructed. Special-purpose regulators based on micros have been designed.

Despite the increasing use of automatic and computer control to improve the operation and safety of process plant, many studies show that automatic equipments are prone to various failures for one or several reasons. One reason why control and protective systems fail to operate properly is that they pay insufficient regard to human nature (Kletz, 1980). It has been reported that a number of incidents have occurred on chemical plants involving computer control systems. These incidents emphasize the fact that computer controlled processes are complex, provide many opportunities for error, and are prone to systematic and random failures (Pearson and Brazendale, 1988).

3.4 Probabilistic Risk Assessment Procedure

3.4.1 Definition - What is PRA?

PRA is an analysis that identifies and delineates initiating events, i.e. possible failures that either directly or through a succession or combination of other events could lead to a major or severe accident such as release of radiation, estimates the frequency or probability of occurrence of each combination, and then estimates the consequences. The probabilities and outcomes of all conceivable consequences are usually calculated on the basis of accumulated data on known failure rates.

3.4.2 Risk Assessment Principles

The fundamental principles of risk assessment are (Cox et al, 1984):

a) that the residual risk (the remaining risk after all proposed improvements have been made) should represent the total risk caused by all possible accidents on the plant;

b) that the spectrum of all accidents should be represented by a finite set whose consequences and expected frequencies should be estimated;

c) that the results should be so presented as to assist designers or decision makers to improve the safety of their plant; and

d) that criteria should be established whereby the results could be judged and decisions made.

3.4.3 Overall Risk Assessment Procedure

The PRA integrates into a uniform methodology (see NUREG-1050, 1984):

- the relevant information about plant design,
- operating practices,
- operating history,
- component reliability,
- human reliability,
- the physical progression of the accident, and
- potential environmental and health effects.

It uses both logic models (such as fault and event trees) and physical models. The logic models depict the combinations of events that could result in a major or severe accident and can be used to determine the frequencies associated with each combination. The physical models depict the progression of the resulting accidents and the damage. The risk associated with any type of accident is the combination (the product) of the frequency of occurrence and the resulting damage. The information extracted from a PRA in the form of predicted frequency of occurrence, resulting damage, and risk provides quantitative and qualitative insights into the aspects of plant design and operation that are the most significant contributor to risk.

The public health effects and economic losses resulting from a major or severe accident can be assessed by means of environmental transport, protective action response, and consequence models. The environmental transport models use site-specific data to predict the spread and fall-out of the released containment. The consequence models use local demographic data to predict the health effects expected to occur in the surrounding population. Throughout the analysis, realistic assumptions and criteria are used. When information is lacking or controversy exists, the individual analysts may introduce conservatisms, increase uncertainties, or evaluate bounds, but the goal of the PRA is to produce an analysis that is as realistic as possible. An integral part of the risk-assessment process should be an uncertainty analysis, which includes not only uncertainties in the data but also uncertainties arising from modelling assumptions.

3.4.4 Problem Definition

Essential preliminaries to the risk assessment study are the explicit definition of the objectives and scope of the study, the clear statement of its purpose and the fundamental assumptions made. It is also essential in any risk assessment to define the plant or system boundaries and collect information regarding its intended design, operation and layout.

3.4.5 Major PRA Steps

Probabilistic or Quantified Risk Analysis and Assessment procedure typically consists of four main stages (see Figure 3.1):

(i) Identification of hazards and failures cases to be considered.
(ii) Evaluation of the consequences of each hazard/failure
(iii) Estimation of the probability/frequency of occurrence of each hazard/failure.
(iv) Comparison of the results of the analysis against specified criteria of acceptability.

(i) Hazard Identification - *What can go wrong?*

The first step in any risk assessment study is the identification of the initiating events (sources of hazards) which could pose significant risks and their causes, for example, equipment failures or operator errors.

Several techniques have been developed for hazard or risk identification. Some of the techniques are: Check-lists, Fault and Event Trees, Failure Modes and Effect Analysis (FMEA) and Hazard and Operability Studies (HAZOP). For a description and

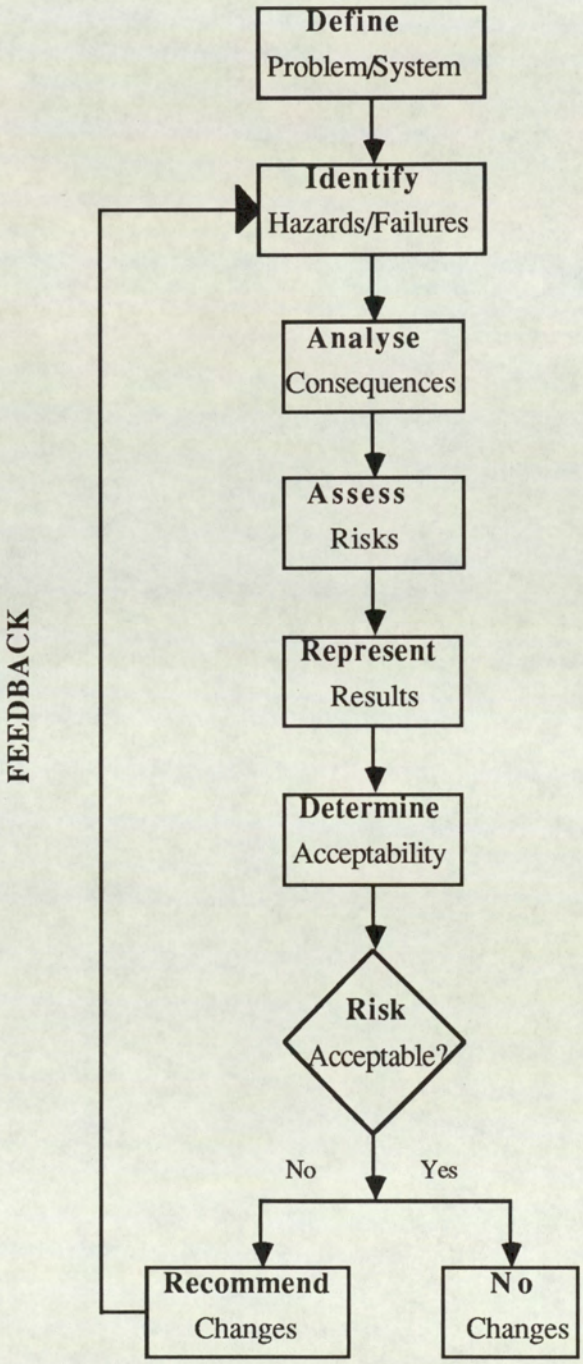


Figure 3.1 Overall Risk Assessment Procedure

discussion of these techniques, see the booklet "*A Guide to Hazard and Operability Studies*" published by the UK Chemical Industries Association (1977).

(ii) Consequence Analysis - *What are the effects & consequences?*

The objective of this process is to determine what will be the potential effects and consequences of each fault/event in terms of degree of damage or injury. This process involves the use of Event Trees to find the various possible outcomes of a given initiating event and the use of mathematical models to determine the relationship between probability or degree of damage and distance

(iii) Risk Assessment - *How often will it happen?*

Having identified hazards/failures and determined their effects/consequences, it is usually important to conduct a quantified or probabilistic risk assessment of these events to identify the factors that pose the highest risk and have the greatest potential for risk reduction. Risk assessment is the process of allocating numerical values to the hazards, faults or events identified to determine their probability or frequency of occurrence and then estimate their consequences. These events include both equipment failures and human errors. If operator recovery from an event within some specified time is considered, the probability of this recovery can be included in the PRA.

(iv) Risk Evaluation - *Is the Risk Acceptable?*

Once the consequences and magnitude of the events identified have been determined and their probabilities of occurrence estimated it is important to determine whether the level of risk to personnel or to the public at large is acceptable and no action will be taken or unacceptable and would warrant some corrective action to reduce it. This process involves comparison of the results of the risk assessment study against specified criteria of acceptability or risk targets. It is also possible to compare the level

of risk with existing code of practice, or existing situations in similar industries, in case the risk objectives are not specified.

3.4.6 Applications of PRA

Over 30 years of nuclear reactor development, the study of risk assessment in many industries has made great strides. Since the US Nuclear Regulatory Commission's publication of the Reactor Safety Study (WASH-1400) in 1975, the use of PRA has increased dramatically. Approximately 30 PRA studies have been or are being performed in the US alone.

A number of companies in process industries other than the nuclear industry, that have in the past relied mainly on the HAZOP technique, have begun to make significant use of PRA in the last decade, as part of their decision-making process on design safety issues (e.g., Canvey).

The many studies that have already been completed are of varying scope. Some of the PRA uses are:

(a) assessment of core-damage sequences.
(b) evaluation of the containment response.
(c) assessment of public risk.
(d) studies of specific accident sequences, such as anticipated transient without scram (ATWS).

Although the purposes of these assessments varied considerably, each study had one or more of the following objectives or end uses in mind:

(a) Identification and assessment of dominant contributors to risk.

(b) Assessment of the plant-specific importance of TMI- and Chernobyl-related requirements and issues.

(c) Assessment of risks at sites with high population densities.

(d) Assessment of specific generic safety issues.

(e) Training of plant personnel.

(f) Development and integration of PRA methodology.

(g) Training in the performance of PRAs.

(h) Assignment of priorities in the use of resources.

(i) Assessment of operating experience and events.

(j) Improvement of operating, testing, and maintenance procedures.

(k) Development of technical information to support recommendations on siting criteria.

(l) Evaluation of emergency-response procedures.

3.4.7 PRA and Human Reliability Analysis

As discussed above, PRA is an approach which has been extensively applied in recent years to the nuclear, chemical, off-shore oil drilling, and other industries to identify the potential risks in a system and to evaluate their probability of occurrence and the expected consequences. Originally, PRA was primarily concerned with failures of hardware components. The Reactor Safety Study (RSS), WASH-1400, published in October 1975, was the first major study to combine hardware reliability with human reliability. Human errors were delineated in the RSS by means of the human reliability analysis (HRA) technique known as THERP (Technique for Human error Rate Prediction) (Swain and Guttmann, 1983). Since then, there has been a growing realization that human actions can have a significant effect on the likelihood of failure of the system. Usually HRA is carried out as part of PRA. The major application of HRA within PRA is to identify the human errors which have a significant effect on the overall safety/reliability of the system (a procedure known as modelling or qualitative

analysis) and to quantify the probability of their occurrence. The main result is a set of Human Error Probabilities (HEPs). In most cases these HEPs are included in the total PRA. The assumptions on which this set of estimates is based are also presented to the system analysts.

3.4.8 Some Limitations and Benefits of PRA

PRA technique has been subject to many criticisms. Some of the major limitations of PRA are the following:

i) Completeness in the analysis - Only those risks that have been identified by using hazard identification techniques such as HAZOP or fault tree analysis (FTA) can be quantified, both these techniques require the exercise of judgement by the analyst. A common criticism of risk assessments based on FTA as a hazard identification is that omissions may occur in constructing the tree and there is, therefore, no guarantee that all possible accidents have been included. However, there is also no guarantee of complete hazard identification using HAZOP method. Not only many deviations may not be recognized as being hazardous, but also the method is not usually applied to all possible combinations of deviations.

The above discussion suggests that the hazard identification will only be as good as the experience, knowledge, and intuition of those performing it. However, this does not mean to invalidate the concept; rather this suggests that a good, solid team be utilized instead of one or two individuals.

ii) Validity of the assumptions made and the data used - The PRA suffers from the same substantial uncertainties as do deterministic analyses, but attempts to address them more explicitly, adds discipline to the evaluation of the operation of a plant, and results in a more complete understanding of risk-important systems and

functions,.interactions among systems, and the importance of human actions. Uncertainty must be considered carefully before a decision is reached. The fact that PRA provides a mechanism to display areas of uncertainty (more so than do conventional deterministic analyses) is actually a strength of PRA rather than a weakness. The weakness that must be guarded against is the tendency to take the PRA point estimates as a given.

Although there will always be uncertainty there is some basis for taking decisions concerning the significance of faults or errors once identified (Holden, 1985). It is important to acknowledge that there will never be completely satisfactory probability data. This is why a probabilistic approach must be adopted which involves a detailed breakdown of each failure into minor events to which estimated probabilities are assigned.

Only actual assessment can reflect the idiosyncrasies of one plant compared to industry averages. Therefore, risk assessment is at least accurate as any other approach. Since the most important contributor to uncertainty is events that are not identified, additional effort should be spent on hazard identification rather than minor improvements to databases.

iii) Possibility of common cause failures - Of particular importance here is human error. Human actions are important in the operation, control, maintenance, and testing of equipment in practically any industrial activity. While beneficial, human actions also contribute to the accident frequency. Past PRA studies have found that beneficial and detrimental contributions of human actions impact the ordering of dominant sequences and the risk of the plant. The studies have shown that human actions can result in the unavailability of plant system before an initiating event, or can cause an initiating event to occur. Beneficial actions include the diagnosis of the nature of an accident and

recovery from an accident sequence. Clearly, PRA technique provide a framework for assessing the importance of human actions in a spectrum of accident sequences.

The definition of specific accident sequences in PRA studies offers the analysts the opportunity to investigate how human actions affect the estimates of accident frequency or risk. For example, the uncertainties in the quantitative impact can be assessed, the ways that operators affect the course of an accident can be described, and the importance of human actions in a particular sequence can be quantified.

The basic human reliability analysis technique (HRA) was first developed in the Reactor Safety Study (RSS), (WASH-1400, 1975). This technique has since been refined and formalized to improve the understanding of the human effect on plant safety. More techniques have also been developed. Hence, inclusion of HRA techniques in PRA studies are undergoing rapid improvement. However, HRA techniques have some shortcomings which will be discussed in some detail in the next chapters. In general the main criticisms have been the non-treatment of cognitive errors, the uncertainties associated with the data base and the complexity and laboriousness of the techniques.

iv) A rather different type of criticism comes from those who regard the whole philosophy of quantification with suspicion. Those who take this view often have little difficulty finding examples of quantitative studies which are open to serious criticism, but, they sometimes tend to impute to the proponents of quantitative assessment claims which the more experienced practitioners are normally careful not to make.

Despite the above cited problems, quantitative risk assessment technology has proven to be an important method for improving the safety of plant design and operation and is now a successful tool and will be increasingly used in technical assessment and

decision making, particularly in the non-nuclear industries where its use has begun most recently.

Among the benefits that accrue from making an integrated and comprehensive quantified assessment of risk are:

1) A comprehensive assessment requires that every aspect of the composition and operation of the whole system or plant is examined in a *systematic way* that identifies the environment the system is exposed to, the interaction between the various components and human operators.
2) Essentially quantitative risk assessment provides an objective *third-party review* of plant design and operating procedures. Such a review assures not only the plant management, but also the nearby community that safety has been given the highest priority in the plant design.
3) The best use of quantified risk assessment is in the *comparison of options* such as alternative processes, alternative sites or alternative safety measures.
4) Quantifying the risks associated with a plant gives the owner an indication of the extent of his potential *financial liability* for compensation for damage resulting from a fault with the plant. This can be useful in calculating the amount of *insurance* cover that is required.
5) There is a concensus amongst practitioners of PRA that the value of carrying out risk assessments goes beyond the provision of *numerical estimates* which provide an input to decision making. *Numerous insights*, both generic and plant specific, particularly, what is important and what is not, can be determined through the use of a PRA study. These insights can lead to modifications to plant design and improved safety and reliability of the plant as a whole. Also, the insights gained are generally not attainable by any other means.

6) The contribution of risk assessment to safe operation stems mainly from the discipline enforced by the need to ask searching questions and obtain detailed information and understanding in order to carry out the analysis. More often than not, this reveals weaknesses or the need for further study or for more information to be obtained. The risk assessment approach provides *a structure on which to base safety programmes*, clarifies and orders the available information and helps to ensure that relevant knowledge and experience is brought to bear.

7) Publication of risk assessments has provided a *common language for dialogue* between the various parties involved in estimation of risk to the public, which has proved to be useful for public inquiries. The most useful feature is often the avoidance of an insistence on absolute safety. Having recognized that absolute safety cannot be guarantied, the acceptability of a situation becomes a matter of the degree of risk. There is little alternative to attempting to evaluate this risk if one wishes to demonstrate that the controls adopted are appropriate.

8) The process of quantifying risks shows where the system can be modified to *improve reliability* and *efficiency*.

9) Quantification of risk gives the regulator *a useful basis for assessing acceptability*.

3.5 CONCLUSION

The main conclusions that can be drawn from the discussions presented above are that measures to assure process plants safety developed historically on the basis of a set of deterministically-defined criteria, which formed the basis for plant design requirements and operational policy. These criteria include the principles of redundancy diversity and the requirement that important systems should be "fail-safe". It was considered that the use of deterministic design principles ruledout releases greater than those assumed to be likely to occur.

Although the safety record of the nuclear industry compares favourably with that of other process industries, the combination of hardware failures, design weaknesses and human errors which have led to accidents such as those at TMI and Chernobyl show that there is a need to extend the scope of safety analyses further than the deterministic approach makes possible.

The rapid growth of the chemical and process industries has occurred in a number of phases, only the later of which has given rise to major industrial hazards in the form of fires, explosions and toxic releases.

Much use of automatic and computer control has been and is being made to improve the operation and safety of process plants. However, automatic equipments are prone to various failures for one or several reasons. One reason why control and protective systems fail to operate properly is that they pay insufficient regard to human nature.

The increased concern and interest in industrial safety expressed by the general public is largely justified by the major disasters which have taken place in the recent past causing many deaths, injuries, and property and environment damages. Similar accidents will happen again, but their probability and consequences can be minimized. One difficulty is that the public often seek guarantees of absolute safety. However, few industries are completely free from danger.

Despite the similarity between the nuclear and chemical industries in respect of safety and environmental matters and in respect of public concern, the techniques adopted for the identification of hazards and assessment of risks are different. The chemical and process industries have adopted the HAZOP techniques for the identification of hazards and design review of their process plants, whilst the nuclear industry has

adopted the technique of PRA. However, the PRA technique has now been more widely adopted throughout the process control industries.

Despite its proven value, the PRA technique has been subject to criticism, such as completeness in the identification and analysis of risks, validity of the assumptions made and the data used, and possibility of common cause failures (including human errors.)

Despite the above cited criticisms, quantitative or probabilistic risk assessment technology has proven to be an important and successful method for improving the safety of plant design and operation and will be increasingly used in technical assessment and decision making, in the nuclear as well as non-nuclear industries.

Clearly, PRA technique generates useful information and insights regarding the design and operation of a process plant, which can be useful to the designer, the manager and regulator in the decision process by providing an improved understanding of the full range of accident sequences and their relative importance. It has, therefore, a vital role to play in the prevention of major accident hazards in nuclear and process industries.

The next chapter examines the safety legislations and quantification of risk and human reliability practices in nuclear and major hazard industries in the UK.

CHAPTER 4

UK APPROACH TO RISK AND HUMAN RELIABILITY ASSESSMENT

4.1 Introduction

As technologies develop and the scale of production becomes larger, in order to reduce overall costs per unit of production, potential danger is generally increased and safety assumes a major role in deciding what type of plant to build and how to build it.

New attitudes, methodologies and legislations have developed with this intensified awareness of safety. Legislation of safety represents the reaction of government to circumstances and to public opinion as modified by consultation with some of the interested parties. It is backed by appropriate bodies whose task is inspection, enforcement, advice, and policy development. Legislation and regulations bring all companies in a country to a minimum accepted level of safety.

Whereas, in the nuclear industry, the TMI-2 accident was the major impetus to the review of the safety approaches then adopted, the interest in the analysis, control and management of major accident hazards was accelerated by the accidents at Flixborough, UK in 1974, and at Seveso, Italy in 1976. These events were eventually followed by the EEC's "Seveso" Directive of 1982 on Major-Accidents Hazards (82/501/EEC). The objective of the Directive was to prevent major accidents and to limit their consequences. The Directive requires Member States to adopt procedures that require manufacturers to prove to the competent authority that they have identified major-accident hazards and adopted all measures necessary to reduce the hazards and limit their consequences.

Although Member States of the EC are required to conform to Directive standards, there is, currently, some variation in approach and response. A particular example of variety of approach is that of the use of quantification in hazard and risk assessments. (see for example, the *"Risk Assessment for Hazardous Installations"*, report published by J.C.Consultancy Ltd., (1986) for the EEC.)

In this chapter the safety legislations and quantification of risk and human reliability practices in nuclear and major hazard industries in the UK are examined. From this examination a number of conclusions are drawn about the main features of the strategies adopted. The legal requirements for presentation of quantified risk assessment as part of the licensing procedure are also discussed, and the ways in which the requirements have been developed for practical application are then examined.

4.2 UK Approach to safety in Nuclear Industry

4.2.1 Overview

In the UK, nuclear installations have to be licensed. The licensing and regulatory process is performed by H.M. Nuclear Installations Inspectorate, which was formed under powers conferred by the Nuclear Installations (Licensing and Insurance) Act 1959. The 1959 Act was modified by the 1965 Nuclear Installations Act and in 1974 the Health and Safety at Work Act was passed and when this act was implemented, Her Majesty's Nuclear Installations Inspectorate (H.M. NII) became part of the newly formed Health and Safety Executive (HSE).

The 1965 Act imposes an absolute liability upon the operators of commercial reactors for any injury or damage caused by the release of radioactive material from their installations and stipulates that no site, except those of the UK Atomic Energy

Authority (UKAEA) and Government Departments, may be used for the purpose of installing or operating any nuclear installation in the UK unless a licence has been granted by the HSE.

NII does not set fixed regulations for design and operation, but rather requires each potential or existing licensee to make their own safety case. The safety case is assessed and negotiated by NII inspectors, and a licence is granted for a specific new installation. Operational reactors and other nuclear plant have to be shut down for regular annual or biennial maintenance, and require NII consent to start up again.

4.2.2 UK Approach to Risk Quantification in Nuclear Industry

At first, the UK NII adopted an approach to safety assessment based on the concept of a *'maximum credible'* accident. In this approach, a worst possible accident was proposed and the plant was designed to accommodate or minimize the effects of the accident,.i.e. if the accident occurs, any consequential release of radioactivity should not cause significant harm to the public. The difficulty with this approach is that it presupposes that any more severe accident is *'incredible'*.(Thomson, 1987) Such a concept of acceptability presents difficulties in application, particularly on how to decide whether theoretically possible failures are credible or incredible and is not useful as a criterion for safety design (The Royal Commission on Environmental Pollution, 1976).

Farmer (1967), suggested a more rigorous approach to the assessment of nuclear plant safety, using probability. The essence of this proposal was as follows: for any given factory or other industrial installation, the acceptable frequency of accidents which may harm third parties varies inversely with the magnitude of the consequences of those accidents. Farmer therefore proposed that nuclear power stations should have to meet a safety criterion proposed in terms of probability and consequence.

Following the publication of the Farmer risk criteria, there was, in official circles, interest in quantified risk assessment which was further stimulated by the publication of WASH-1400 in 1975.

The fundamental principle currently applied in the UK to the regulation of all industrial risks is the so-called ALARP (as low as reasonably practicable) principle, which is the keystone of the UK's Health and Safety at Work Act (HASAWA) legislation. This requires that the operators do whatever is reasonably practicable to reduce risk, bearing in mind the costs of further reduction.

Although the onus is on the licensee to make his own case, the NII has set out its safety assessment principles and provided detailed guidance on how to use the ALARP principle for safety assessment of power reactors (HSE, 1979a) and of nuclear chemical plants (HSE, 1983a). These guidelines described the fault sequence evaluation which the Inspectorate would expect to see for any nuclear installation proposals and postulated the main quantitative conditions that should be satisfied under normal operation and fault conditions. If the risks are greater than specified in the guidelines, then some improvement will need to be made to reduce the risk to an acceptable level.

The principles distinguish between accidents which give rise to dose equivalents received by the public of no more than one Emergency Reference Level (an ERL is the radiation dose below which counter measures are unlikely to be justified) and those accidents which might give rise to larger doses.

In the 1983a HSE's document, Section 3.10 *"Analysis of Plant Faults, Transients and Abnormal Conditions"*, the general aim of the *"fault analysis"* has been described as follows:

"to predict, when reasonably practicable, the behaviour of the plant and associated equipment in specific fault conditions, and to estimate the consequences of such faults and the likelihood of their occurrence, in quantitative terms"

The reports to the NII published on *"the generic safety issues of PWR reactors"* (PWR, HSE, 1979) and on the review of the Sizewell B pre-construction safety report (Sizewell B, 1982) illustrate the role that quantification of risk in probability terms is expected to play in future assessments.

The HSE's new discussion document, *"The Tolerability of Risk from Nuclear Power Station"* (HSE, 1987), is a first step to fulfilling a recommendation made by Sir Frank Layfield in his report (1982) on the Sizewell B Public Inquiry that the HSE should formulate and publish guidelines on the tolerable levels of individual and societal risk to workers and the public from nuclear stations (see Table 4.1 for new HSE's risk objectives).

Table 4.1 UK HSE's Risk Levels for NPPs (1988)

Type of risk	Probability per reactor year
1.Limiting design basis accident (DBA).	1×10^{-4} to 3×10^{-4}
2.Any "uncontrolled" (beyond DBA) release at each reactor.	1×10^{-5} to 1×10^{-6}
3.Any uncontrolled release large enough to produce doses of 100mSv (the ERL dose) within 3km.	Near to 1×10^{-6}

The authors of the HSE's document on tolerability of risk say that:

> "(It) *breaks new ground in setting out in detail the basis for HSE's assessments of the civil nuclear risk, and their approach as licensing authority to its control.*"

The document discusses tolerable levels of individual and societal risk. With regard to individual it concludes that, for the public, risks of death higher than 10^{-4} per year are intolerable, whereas risks less than 10^{-6} per year might be broadly acceptable. For societal risk it suggests that a significant nuclear accident might be accepted as just tolerable at a frequency of 10^{-4} per year.

4.2.2 UK Approach to Human Reliability in Nuclear Industry

The HSE's approach to human reliability and human error has been described by the NII's human factors specialist, Whitfield in two recently published papers:

- *"Human Reliability from a Nuclear Regulatory Viewpoint"* (Whitfield, 1987a)
- *"A Regulatory Perspective on Human Factors in Nuclear Power"* (Whitfield, 1987b)

Whitfield (1987b) described broadly the role of NII as consisting of assessment of design proposals for new plant and modification, and inspection of operating installations. He said that in both these areas, human factors (or ergonomics) is considered under a broad range of headings:

(a) System design

(b) Operator-plant interface

(c) Operating procedures

(d) Operating environment

(e) Selection and training of operating staff

(f) Organization and management.

In assessment, each of these aspects is reviewed at the appropriate stage of the presentation of the safety case. In inspection, possible conditions for human error are reviewed.

Whitfield (1987a) described the NII approach to human error and human reliability as below:

i) NII and Human Error

Whitfield states that,

> *"Any evaluation of safety must take account of human error, and NII has recognized the importance of ergonomics for some times."*

NII has convened a working group on ergonomics/human factors issues (HSE, 1982) and ergonomics is specifically dealt with in part of the NII review of the Central Electricity Generating Board's pre-construction safety report (HSE, 1983b). The major topics identified there were:

(a) Design and construction:

(b) Operational management:

(c) Fault studies:

An extensive programme of ergonomics development is now being pursued by CEGB, and this will be subject to the further NII assessment in the Sizewell 'B' design. The identification of possible human errors in operation and maintenance, and their reduction and mitigation in accordance with design safety goals, is an important component in that work.

Ergonomics issues were included in the recent NII audit of part of the Sellafield nuclear reprocessing plant, where specific attention was given to control room design, operating procedures, and plant operations (HSE, 1986).

ii) NII and Human Reliability

The NII procedure to establish safety targets and to minimize the dangerous consequences of human fallibility have been summarized by Whitefield as follows:

> *"The safety targets for a station like Sizewell 'B' are that the frequency of a limited release of radioactivity to the environment should not exceed 1 in 10,000 per reactor year, and that the frequency of a large uncontrolled release should not exceed 1 in 1,000,000 per reactor year. The licensee constructs fault and event trees, to demonstrate that significant initiating events, and likely combinations thereof, do not breach these limits. The current policy of NII is that comprehensive quantification of human reliability is not feasible; nevertheless, it is necessary to show that all required human safety actions are within normal operator capabilities, and to have as much evidence as possible that operator error, passive or active, will not seriously qualify the outcome of the fault analysis."*

> *"Of course, it is very doubtful that human error rates per se are compatible with the frequency goals set out above, particularly taking into account the varieties of human involvement in...activities of the station. There has to be considerable dependence on highly reliable automatic systems, particularly for preserving reactor safety in abnormal conditions or after a major fault has developed. In addition, principle 124 (HSE, 1979) requires that the safety case should not rely on any operator intervention for 30 min. after a major fault or reactor trip (automatic shut-down). The operator is expected to monitor the series of operations of the safety systems during this period, and to reinforce the series if required. Obviously, this implies exhaustive analysis, at the design stage, of his possible errors or misperceptions. It may be necessary to prevent any operator actions which threaten the success of the automatic systems. Progress is being made in analysing and solving such problems."*

Whitfield (1987b) stated that,

> *"NII does not considers it practicable to require comprehensive quantification of human reliability in the same manner as the rest of the systems."*

He emphasized, however, that:

> *"it is necessary that human error be significant or comparable with the target."*

He added that a design and assessment strategy acceptable to NII would be:

(1) Identification of the major cases where human error could make a significant contribution to the probability of a major release.

(2) Analysis of each of those areas separately, to show that the task requirements are within human capabilities, and that, by judgement or quantification where feasible, the probability of human error leading to a major release will not exceed the target.

(3) Particularly where high reliability is required, the analysis must include consideration of:

   (a) the time available to correct human action, and the operator workload involved;
   (b) arrangements for independent checking of human action;
   (c) the contribution of selection and training practices, design of procedures and interfaces, and organizational features, to correct human action.

4.3 UK Approach to Safety in Major Hazard Industries

4.3.1 Overview

In the UK legislation on major hazards is directed towards making industry as safe as practicable means can provide, and separating the public from residual potential hazards by control of land use and of building development.

In 1972 the recommendations of the Report on Safety and Health at Work produced in 1972 by a committee chaired by Lord Robens resulted in (HMSO, 1972):

(a) The introduction of one *Act:* the Health and Safety at Work etc. Act (HASAWA) in 1974. The Act is based on *self-regulation* and imposes general (and in some areas specific) duties on all people at work. Everyone is required under these duties to prove that he has taken all reasonably/ best practicable measures to ensure safety. A summary of duties under HASAWA is shown in Table 4.2.

(b) The formation of one *Authority*: the Health and Safety Commission (HSC), to co-ordinate health and safety at work. The HSC, which consists of representatives from both sides of industry and local authorities, decides policy issues and instructs its executive body, the Health and Safety Executive (HSE) to act on its behalf.

The establishment of the *Advisory Committee on Major Hazards* (ACMH) by the Health and Safety Commission towards the end of 1974 was partly a result of the Flixborough incident in June 1974. The committee was asked (HSE, 1976):

Table 4.2 Duties under HASAWA (From Carson and Mumford, 1988)

Section	Duty
Employers	
2 (3)	To prepare and up-date as often as necessary a written statement of his general policy with respect to health and safety at work of his employees. This must also describe the organization and arrangements for implementing the policy, and it must be brought to the attention of all employees.
2 (1)	To ensure the health, safety and welfare of all employees.
2 (2) (a)	To provide and maintain safe plant and systems of work.
2 (2) (b)	To ensure the safety and absence of risk in connection with the use, handling, storage and transportation of articles and substances.
2 (2) (c)	To provide the necessary information, instruction, training, and supervision for the health and safety of employees.
2 (2) (d)	To ensure the place of work is maintained in a safe condition with safe access and egress.
2 (2) (e)	To provide and maintain a safe working environment which is without risk to health and adequate for their welfare at work.
2 (6)	To consult with safety representatives
2 (7)	To set up a safety committee
3	To conduct his undertaking so as not to expose the general public to risks.
4	To prevent those, who are on his premises but who are not his employees (e.g. visitors, contractors), from being exposed to risks in their health and safety.
5	To use best practicable means for preventing noxious or offensive dusts or fumes from being exhausted into the atmosphere, or that are harmless.
Employees	
7 (a)	To take reasonable care of his own safety and health and of those who may be affected by his acts or omissions.
(b)	To co-operate with the employer in discharging his duties.
8	Not to interfere or misuse anything provided for health safety or welfare reasons.
Designers, manufacturing importers or suppliers	
6	To ensure their products are designed and installed so as to be safe when used and to test and provide necessary data on the product for safe usage.

"To identify types of installations (excluding nuclear installations) which have the potential to present major hazards to employees or to the public or the environment, and to advise on measures of control, appropriate to the nature and degree of hazard, over the establishment , siting, layout, design, operation, maintenance and development of such installations, as well as over all development, both industrial and non-industrial, in the vicinity of such installations."

The ACMH'recommendations (HSE, 1976, 1979b and 1984) and the EEC Directive provided the basis for, the *Notification of Installations Handling Hazardous Substances* (NIHHS) Regulations (1982) which were implemented in January 1983, and the *Control of Industrial Major Accidents Hazard* (CIMAH) Regulations (1984).

The CIMAH came into force in January and April 1985 to implement the EEC's Seveso Directive. Major legislation relating to industrial safety in the UK is indicated in Table 4.3.

Table 4.3 A Sample of Industrial Safety Legislation in the U.K.

1831	-The Factories Act
1906	-The Alkali Act
1927	-The Explosives Substances Act
1961	-The Factories Act
1972	-The Robens Report
1974	-The Health and Safety at Work etc. Act . (HASAWA)
1974	-The Advisory Committee on Major Hazards reported in 1976-79-84 (ACMH)
1982	-The Notification of Installation Handling Hazardous Substances Regulations (NIHHS)
1984	-The Control of Industrial Major Accident Hazard Regulations (CIMAH)

4.3.2 UK Overall Approach to Major Hazards Control

The UK approach to the control of major industrial hazards, derived from the recommendations of the ACMH and confirming European requirements is centred in the following principles or stages (see Table 4.4):

Table 4.4 The UK Approach to Control of Major Industrial Hazards

CONTROL STEPS	REGULATIONS
Identification	via NIHHS 1982, CIMAH 1984, 1988
Risk Reduction	via HASAWA 1974 and CIMAH
Mitigation	via CIMAH (emergency planning and information to the public) and land-use planning control

1. Risk Identification

The NIHHS require operators of site to inform the HSE if more than a specified minimum quantity of a defined hazardous substance is stored, manufactured, processed or used at an installation.

2. Risk Reduction

The main objective of risk reduction, which involves risk assessment and appropriate control on site,.is to reduce the risk of an accident to a low level, but it does not eliminate the hazard completely. There remains some residual risk. The *general* requirements of the CIMAH Regulations apply to sites which store or use hazardous substances. In such cases the operator of the site must:

a)- notify the HSE of any major accident which has occurred on his site, with details taken to prevent its recurrence;

b)- be prepared to demonstrate to an Inspector, on request, (and produce documentary evidence as appropriate) that he has considered the potential for major accidents from his operations, and he has taken all appropriate measures both to prevent their occurrence and to limit the consequences of any which may occur.

Additional, more *specific* duties apply to installations on which are stored or used certain substances in excess of specified thresholds. Operators of such installations are required to:

a)- prepare (and keep up to date) on- and off-site emergency plans.

b)- provide appropriate information to the public.

c)- submit (and keep up to date) a written *"safety case"* to the HSE, at least three months before commencing any new activity, or in the case of an existing activity before 8 January 1989.

The safety case should:

a)- identify the nature and scale of use of dangerous substances.

b)- place the installation in its geographical and social context.

c)- identify the type, consequences, and relative likelihood of potential major accidents.

d)- identify the control regimes and systems on the site, and in so doing demonstrate that the operator of the site has considered and (presumably) is satisfied with the adequacy of his controls, and that any residual risks are at an acceptable level.

3. Accidents Mitigation

The main objective here is that the element of residual risk is taken into account in planning decisions. It is concluded in the ACMH that the responsibility for such control should remain with land-use planning authorities. HSE's role is to offer advice to the planning authorities about the risks associated with major hazards, and the potential effects upon populations in their vicinity. Where HSE's advice is to refuse planning permission, it is accompanied by an offer to supplement the general advice with more detailed advice based on the assessment of risk associated with the hazardous installation.

4.3.3 UK Approach to Quantification in Hazardous Installations

Having outlined the organization that has been established to deal with the regulation of major hazards, the way risk acceptability is judged in practice can now be examined.

The original approach used by HSE to advise planning authorities about the risks associated with major hazards was based on a concept of *'protection'* of those exposed to a hazard. The application of this concept involves the identification of the worst events (of fire, explosion, or toxic release) and then the determination of a separation distance based on a defined level of injury or the intensity of the thermal radiation, blast overpressure or toxic exposure respectively.

The protection concept has been subject to criticism on a number of grounds; these include (HSE, 1989b):

(a) the possibility that the protection provided is beyond that which is 'reasonable', if a low probability of serious injury is combined with a very low likelihood of the critical event, thus resulting in excessive restrictions on land use;

(b) the somewhat arbitrary nature of the worst event, and potential inconsistency between installations in deciding which major event to use as a basis;

(c) the difficulty of comparing the degree of protection with that which seems to be necessary or desirable for other hazards in life.

There has been no official requirement laid down stating in uniquivocal terms the quantitative risk criteria that an installation has to satisfy to be acceptable. However, the first report of the ACMH does give some indication of the views of proximate policy makers on the acceptable probability of serious accidents:

> *"If, for instance, such tentative conclusions indicated with reasonable confidence that in a particular plant a serious accident was unlikely to occur more than once in 10,000 years (...), this might perhaps be regarded as just on the borderline of acceptability..."*

The theme of quantification of risk was carried through to the third and final report of the ACMH in the following way:

> *"We hold the view that the hazard surveys we recommended in our previous reports would in part be based upon some form of quantitative assessment. We note that the discipline of quantitative assessment is still developing rapidly and is only one useful element in the overall system of control and only one input to be considered in the decision making on major hazard plants."*

> *"We believe the HSE should develop and make known the guidelines of what is good current practice."*

The fact that there were no stated quantitative criteria for acceptability that have to be satisfied has made the quantitative risk studies that were carried out to determine the acceptability of expansion of the chemical complex on Canvey Island more interesting. The HSE was requested to carry out an evaluation of the risks to the public. The HSE

asked the Safety and Reliability Directorate of the UKAEA to evaluate the risks associated with this proposal. The Canvey report was published in 1978 and reassessed in 1981. The results of the assessment of the risks were presented in a way that demonstrated in quantitative terms the benefits in terms of reduced risk that could be obtained by modifications that could be made to various installations on the island.

This study clearly showed that it was possible, albeit with some uncertainty and at significant expense, to estimate this kind of risk. The value of such a study was made clear by the inspector of the most recent public inquiry into the installations at Canvey, who recorded (1982) the agreement of the major parties represented that risk assessment was an essential first step in taking decisions about the safety of such installations.

Increasing use has been made of numerical techniques of risk assessment in the process of decision making about safety of major hazard installations, and the HSE considers that it is now time that its advisory role be "risk-based" and for criteria to be established and made widely available for discussion.

The quantified risk criteria which the HSE has developed to provide advice to local authorities about proposed land-use developments in the vicinity of major industrial hazard installations are set out in its recently published booklet, *"Risk criteria for land-use planning in the vicinity of major industrial hazards"* (HSE, 1989b).
The document states that :

> *"HSE's advice should take specific account of the likelihood of injury to the public, as well as the possible extent of injury effects. The protection concept takes account of likelihood in a qualitative way, but it seemed to be expected that it should be quantified if possible."*

To meet this need, HSE uses quantitative or probabilistic techniques (QRA or PRA), similar to those described here in Chapter Two, to quantify the risks associated with hazardous installations, in order to uderpin its advise to planning authorities. However, HSE's emphasizes that there may still be a need for particular cases to be judged beyond the strict application of the criteria and cases where the quantification of probabilities is so difficult or uncertain, or potentially misleading, that HSE's advice continues to be based on the *'protection concept'*.

The document discusses the implications of risk assessment, considers both individual and societal risk and suggests criteria which may be appropriate.

For new developments in the vicinity of existing major hazard installations, HSE suggests the following criteria:

i) Individual risk

(a) Lower bound - For a typical individual in a group at risk, HSE will use the figure of:

> *"1 in a million per year, in relation to the risk of receiving a 'dangerous' dose or worse, for a typical pattern of user behaviour in a development."*

(b) Upper bound - HSE will use an upper limit of

> *10 in a million per year of a dangerous dose or worse,* for all development cases above a certain size (see HSE, 1989, para 57).

ii) Societal Risk:- *A Judgemental Approach*

HSE notes that it is difficult to define the size of a disaster risk in terms of numbers. With societal risk, unlike individual risk, it is not only the chance which is important

but also the potential size of the disaster. However, there is at present no clear consensus on criteria for societal risk and it is not even clear how best to describe such risk. Therefore, rather than attempt to produce numerical criteria for societal risk, HSE's advice will be based on the criteria for individual risk. The element of societal risk will be allowed for by using a harsher judgement for larger developments in the middle zone of the criteria. The judgement here is essentially a weighting of the significance of individual risk to allow for the larger size of development, on the basis of a judgement of the aversion of society to larger-scale disasters.

4.3.4 UK Approach to Human Error in Hazardous Installations

HSE (1989b) considers that human actions are one of the factors of uncertainty involved in the estimates of probabilities of risk that should be taken into account. Human actions/errors influence all aspects or stages of safety control of hazardous installations. HSE believes that automating the parts of a system which are critical for safety is not alone enough to reduce human errors. Usually there will be a mixture of automation and human operation in a system. Human error must therefore be,
"taken into account as a cause of accidents, to give a full assessment of risks from an installation."

HSE identifies two ways of assessing the influence of human errors:

- *implicitly* (using data of failure-rates from all causes) or
- *explicitly* (by analysing the potential causes of failure including human error).

The HSE methods rely mainly on the implicit method. This approach is based on the assumption that historical data for component failures include those caused by human error. Then the calculated risk should relate to an "*average*" level of human error. HSE suggests the inclusion of an adjustment to the predicted failure-rates to allow for some deviation from 'average' of the overall quality of the safety management and notes that

an assessment based on 'average' levels should be somewhat pessimistic, since standards should tend to improve.

In addressing the question of *"grossly negligent"* or *"perverse"* human actions which might defeat the best precautions built into a plant, HSE identifies the following possible causes:

1. failures of management,
2. false management goals (production before safety),
3. gross deviation from normal behaviour, or deliberate acts.

HSE states that these deviations *"must be allowed for in judging the predicted risk figures"*. The approach applied here is similar to that adopted for risks from nuclear power stations (see HSE, 1987) and stated in (HSE, 1989b) as below:

> *"it was reasonable to assume that the risk of a release of substance is within an order of magnitude of that assessed on the basis of plant failures from historical data, for an installation operating with strict standards of control, training of staff etc. This allows for all causes of accidents including gross human actions, but it does contain an implicit judgement on the quality of management. For plant that is less well managed the probability of a serious accident could be greater. The way to deal with this is by achieving proper quality of management and operation, rather than to apply large factors to QRA results."*

In order to cope with the effects of uncertainties in hazard and risk assessment, HSE uses an approach described as *"cautious best-estimate"*. It uses realistic, best-estimates assumptions, but where there is difficulty in justifying an assumption, some overestimate is preferred. It is believed that this approach (HSE, 1989b):

> *"helps to offset any uncertainty arising from the possibility of grossly abnormal human behaviour and other unquantified causes of accidents."*

4.4 CONCLUSION

The general conclusions that seem to be justified from the examination of the UK practices in nuclear industry is that quantified risk analysis is used by the regulatory authorities as a way of assessing the acceptability of a particular proposal and that the fundamental principle applied in the UK for severe accidents (which would lead to high off-site doses) is the ALARP principle, which specifies that the risks should be made as low as reasonably practicable. The indications are that the NII use the estimate of reliability in an iterative way to encourage the proposer to progressively improve up to an acceptable but maintained standard. That is, if the reliability of some system or component is indicated as being low in the pre-construction safety report, the proposer will be asked to improve the design of that item so that by the time the final safety report is presented it will have an acceptable level of reliability.

The HSE's document on the tolerability of risk, the procedure employed to establish risk targets and the qualitative approach to human reliability have been subject to many criticisms. Two recent criticisms of interest here were made by Gifford (1989) and Reason (1987).

Speaking at the Hinkley Point C Inquiry and commenting on the HSE's document on risk tolerability, Gifford said that the document had underestimated the risks from nuclear power stations and the probability of serious accidents caused by human error.

> *"The document contains little of the scepticism and reserve that should be essential parts of the philosophy of an enforcement agency",* he added.

Commenting on the HSE.NII's safety targets and its qualitative approach to human reliability, Reason asked: *"How can meaningful safety targets exclude human*

reliability?" He said that whereas the CEGB is required to identify and quantify system failures, the HSE's numerical safety targets take no direct account of human fallibility. However, Reason presented many sources of information about causes of nuclear power plant accidents from the USA and Europe which all indicate that human performance problems at all stages constitute by far the greatest proportion of root causes. He gave as an example the Chernobyl disaster *"at which a group of award-winning operators blew up a comparatively safe reactor without the assistance of a single component failure."* He concluded that such observations do not inspire confidence in safety targets that exclude human factors issues.

The conclusions about the UK's organization in the field of major hazard assessment that seem to be justified are that there is an involvement of the local authorities with the assessment process and some form of safety document has to be presented. Also there are no specific quantitative criteria in terms of probability that have to be satisfied.

The quantified risk criteria which the HSE has developed to provide advice to local authorities on the development of land in the vicinity of major hazard installations have been set out in its most recently published document (HSE, 1989b). This document has also described the HSE approach to human errors. In general, HSE recognizes the importance of human contribution to major hazards, although, as for nuclear industry, this is not stated quantitatively.

The next two chapters discuss the nature of the human reliability analysis (HRA) domain problem. Whereas Chapter 5 discusses the historical and qualitative aspects of HRA, Chapter 6 presents a critical review of major quantitative HRA approaches.

CHAPTER 5

QUALITATIVE APPROACHES TO HRA

5.1 Introduction

To develop the proposed expert system, the details of solving the problem selected for the system must first be understood. Therefore, extensive literature reviews were performed to gather as much as possible information about the problem domain. To gain a better understanding of the expert's problem-solving strategies, a thorough analysis of the human reliability analysis problem domain is necessary.

In recent years increasing attention has been paid to the study of human reliability, particularly in large scale technological systems such as nuclear power plants (NPPs), the chemical and petrochemical industry. Today's systems have become highly sophisticated and complex and involve large aggregations of mass and energy which are subject to centralised control by relatively few individuals. This means that the consequences of human errors or equipment failures can be catastrophic both for the system and for the society at large. The TMI accident was the result of a combination of human errors and hardware failures. Various studies have indicated that a significant proportion of system failures were due to human errors. According to Dougherty et al., (1988), the estimated risk due to human action or inaction in nuclear industry is about 50 to 70% and offshore petroleum drilling hazards apparently arise in some 70% of the cases from human causes.

In addition to safety considerations, a high level of human reliability is also essential in order to maximise the availability and productivity of such systems. It is not sufficient to consider only the operational aspects of human involvement. The impacts of human reliability must be considered during all the system phases form project conception,

through design, manufacture, commissioning, operation, inspection, testing and maintenance to the final stage of decommissioning, as well as future modifications as examplified by the Flixborough disaster 1974.

Nowadays, the effort is directed toward replacing the human functions with machines in order to reduce the occurrence of human errors. However, machines can fail to work and even highly automated systems do not totally remove human involvement.

This human contribution to the system safety, reliability and availability can be understood, assessed, and quantified using techniques of human reliability analysis (HRA).

There are, however, several technical problems in trying to assess human reliability and because HRA is a controversial, fairly new discipline, it has been criticised on a number of points (Adams, 1982, for example). Until relatively recently, the level of theory that has been utilised in HRA has been relatively unsophisticated (Embrey, 1987). This is mainly because there is no theory of human behaviour that is both well-accepted in the various human sciences and complete enough to use to develop a theory of human reliability (Dougherty and Fragola, 1988). However, Rasmussen et al., (1987) point out that HRA is founded on the idea that the seemingly immense variety of human errors observed may reflect complexity of the environment, rather than complexity in the psychological mechanisms involved.

The main purpose of this chapter and the next two chapters is to define the nature of the domain problem of this research project, by identifying the key concepts and their interrelations, the human reliability expert's tasks and subtasks that the proposed expert system is expected to perform, and the aspects of expertise that are essential in their performance.

The emphasis in this chapter is on the historical, theoritical and qualitative or modelling aspects of the HRA procedure. The quantitative analysis of human reliability is dealt with in both Chapters 6 and 7.

5.2 Historical Review of Human Factors/Ergonomics

There is a natural relationship between HRA and human factors/ergonomics disciplines. Both are concerned with analysing, predicting, and improving system performance:

> (HRA) *is an offshoot of the analysis of human performance in an industrial setting, which in turn, is one of the many human factors concerns in industry"*
> (Dougherty and Fragola, 1988).

Human Factors originally developed as a result of military needs. During World War I, the governments of the US and of the UK directed significant attention to military personnel selection and training. The prime target of this effort was *"fitting the man to the job."* In 1918, in the US, laboratories were established at the Wright-Patterson Air Force Base and the Brooks Air Force Base to perform human-factors-related research (Meister et al.,1965). These laboratories have performed research on areas such as complex reaction time, perception and motor behaviour.

The years between the two World Wars witnessed major growth in disciplines such as industrial psychology and industrial engineering. During World War II engineering systems became so complex and sophisticated that the need for human factors consideration became mandatory (Dhillon, 1987).

By 1945 human factors as a specialised discipline was recognised (Dhillon, 1987). At present, it is widely used in aerospace, nuclear and military-defense, industry, and in

many other areas. Several textbooks on the subject have appeared and a number of research journals are devoted to this field.

Human factors approach is of an interdisciplinary nature and has various names: *"Human Factors Engineering"*, *"Human Engineering"*, *"Human Factors"*, and *"Ergonomics"*. The first three terms are used most widely in the United States. The term "Human Engineering" is now falling into disuse by human factors practitioners since some persons have used it in the procrustean sense of engineering humans, i.e., making people fit the environment -- the antithesis of human factors engineering (Swain and Guttmann, 1983). The last term, "Ergonomics", is used most frequently in Europe and in other countries but is now becoming popular in the US as well. The discipline is concerned with designing hardware, software, procedures, information, and work environments *"for the human use"* (McCormick & Senders, 1982) and so that they match human capabilities and limitations (Chapanis, 1965)

5.3 Historical Review of Human Reliability

To understand the genesis of human reliability (HR), it is necessary to place it in its chronological and conceptual context. The roots of HR can probably be traced back to the beginnings of the Industrial Revolution but HR began to evolve into a discipline in the 1950s. Under the auspices of industrial engineering and behavioural psychology, initial efforts were made to investigate the influence of the human being in the performance of tasks.

The development of the field of human reliability analysis was influenced by the first systematic treatements of the problem of reliability in complex technical systems. Equipment reliability engineers had already possessed well established techniques for quantifying the effect of equipment performance on system output, but had ignored the personnel element because they had no means of dealing with it (Meister, 1985). The

successful application of reliability engineering techniques to the evaluation of hardware systems motivated human factors specialists (HFs) or ergonomists to attempt to apply the same tools to the evaluation of human performance.

HFs were often organisationally part of the reliability or quality assurance division (Meister, 1985). In this context, many of them felt it necessary to align their thinking with that of the host discipline. They pointed out that numbers assigned to systems were incorrect because they ignored the effect of personnel on the system, thus making reliability estimates overly optimistic. The work of Williams (1958) and of Shapero et al. (1960) was influencial in pointing out the effect of human errors on system performance. Williams was one of the first persons who recognised that human reliability must be included in the system reliability prediction. Shapero et al. pointed out that human error is the cause of a large proportion (i.e., from 20 to 50%) of all equipment failures.

Moreover, in working with design engineers HFs were often challenged to quantify their recommendations and the rational for these recommendations. If a metric describing the performance of the personnel element in the system was to be developed, it was desirable that it be the same as that for equipment elements. This would permit the combination of separate estimates for equipment and personnel into a single system prediction. This consideration has the effect of disposing researchers interested in quantifying human performance to think in terms of adopting the equipment reliability methodology (Meister, 1985).

However at the time, there was very little in the way of human performance data and also no accepted human performance theories or models. The realisation of this led to a research project which produced a prototype demonstrational data base containing human reliability figures known as the AIR (American Institute for Research) Data Store in 1962 (Munger et al., 1962). Around the same time, the pioneering work of

Rook (1962, 1965) and particularly of Swain (1963) and his colleagues at the Sandia National Laboratories began. In 1964, several approaches to quantifying human performance were developed using the AIR DS (*Human Factors*, 1964). A noticeably missing factor in these quantification schemes was a systematic approach to the classification of human performance in various tasks. Classification structures based on behaviour were attempted by several people, among them Berliner (Berliner, 1964). In this same time frame, the human performance model problem was being studied by Swain and his collaborators. His early work (Swain, 1964) was later refined into the well known *Technique for Human Error Rate Prediction* (THERP), work which still continues

In August 1973, the journal entitled Institute of Electrical and Electronics Engineers (IEEE) Transactions on Reliability (Regulinski, 1973) published a special issue devoted to HR. A number of excellent papers appeared in this issue.

The 1975 Reliability and Maintainability Symposium (Proceedings, 1975) exhibited examples of HRA models which had been extended to different applications earlier in the 1970's. These proceedings included a paper on the Siegel et al. efforts to produce *"a set of stochastic, digital simulation models for simulating the performance of the human component in man-machine system..."* In the same proceedings, Swain and Guttmann described the application of THERP to the nuclear power plant environment. This paper was a brief overview of the work that they had performed for the Reactor Safety Study (RSS), WASH-1400 (USNRC, 1975), which was the first major study to combine hardware reliability with human reliability. Also in this time frame, the US Navy published a user's manual for their NAVSEA Human Reliability Prediction System (US DoD, 1977).

Then, the TMI accident occured in March 1979. It was the worst nuclear accident up to that time. Such an accident was largely inconceivable to the engineering community.

Following the TMI accident, many committees were formed and a number of major investigations were conducted to identify the cause(s) and to propose safety improvements (see Hagen and Mays, 1981; Farr, 1984). The report resulting from the President's Commission (Kemeny, 1979) charged with review and assessment of the accident, found that *"inappropriate operator action"* resulting from training and procedural deficiencies, failure to learn from previous incidents, and deficient control room design caused the TMI accident (some people, for example Perrow (1984), have noted the unfairness in attributing the operators with any such blame.)

A weeklong forum/workshop, sponsored by the IEEE, NRC and Brookhaven National laboratory, was convened, 2-7 December 1979 to discuss human factors and nuclear safety. This first "Myrtle-Beach Conference" brought together representatives from engineering, psychology, reactor operation, and HRA. One of the priority research areas identified from Myrtle Beach I was for a *"systematic, consistent, and reproducible approach for the quantitative evaluation of the reliability of the human component in the system"* (Schmall, 1980). An initial draft of the documentation of THERP, in a handbook format, was available at the conference, which included one such approach to HRA. A formal draft version, *"Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Application"*, was published in 1980 and the final version in 1983 (Swain and Guttmann, 1980-83).

Another development in human reliability studies was the increased recognition of the importance of the higher level cognitive functions such as diagnosis and decision making. A number of mental models have been proposed. Dougherty and Fragola (1988), believe that:- *"A solution to the human performance classification problem resulted from a "model" created by Rasmussen."* (Rasmussen, et al., 1981). A simplified form of this model is given in Figure 5.2.

A second Myrtle Beach conference was held in September 1981. Swain's THERP technique, a cognitive-oriented model of Moray, the Operator Action Tree (OAT) model of Wreathall, and an HRA approach of Hannaman were reviewed (Conference Record, 1982). The "ideal" model, it was agreed, would combine the features of all types, by basically enhancing the use of a cognitive model with a top-down system engineering approach. At that time, it was agreed that THERP provided an acceptable model for certain situations but that some cognitive modelling was necessary to provide insight into operator behaviour, especially in off-normal process conditions. Notably, however, the PRA Procedure Guide (NUREG/CR-2300, 1983) adopted THERP as its only recommended HRA approach in 1981.

Ever since, further advances have been made in the HR field and there has been considerable growth of the published literature on the topics of human performance and man-machine systems (see for example, Lee, 1988).

5.4 System Reliability and Human Reliability

Both human reliability and system or engineering reliability are generally stated in terms of the probability of a function/task versus time.

Green and Bourne (1974) define system reliability as:

> *"that characteristic of an item expressed by the probability that it will perform its required function in the desired manner under all relevant conditions and on the occasions or during the time intervals when it is required to perform."*

The traditional usage of the term *Reliability* in psychology is as *consistency* (repeatibility) in performance or judgement (Meister, 1985). Reliability is used here to measure performance *accuracy*, or its converse error.

Reliability is conventionally defined (Evans, 1976) as the probability of successful performance of a mission.

Meister distinguishes between HR as a measure and HR as an activity or HRA.

HR as a measure is (Meister, 1966):

> *"The probability that a job or task will be successfully completed by personnel at any required stage in system operation within a required time (if the time requirement exists)."*

HR as an activity is (Meister, 1985):

> *"The analysis, prediction and evaluation of work oriented (MMS) human performance in quantitative terms using, for example, such indices as error likelihood, probability of task accomplishment, and response time."*

Swain and Guttmann (1983) defined HR as:

> *"The prpbability that a person (1) correctly performs some system-required activity in a required time period (if time is a limiting factor) and (2) performs no extraneous activity that can degrade the system."*

The usual measure of HR that is used is the ratio between the number/frequency of successful completion of a given task (s) and the number/frequency of times the task has been attempted; that is (n), i.e.,

Human Reliability (HR) = s/n

HR expressed in probabilistic form is *1.0* minus the *Human Error Rate*.

Where:

1.0 means invariably correct (successful) performance, and
Human Error Rate is equivalent to task failure.

In the Handbook by Swain and Guttmann (1983), the use of *human error rate* has been avoided, instead the basic index of human reliability used is the *human error probability* (HEP), i.e., the probability that when a given task is performed, an error will occur. An HEP is calculated as the ratio of errors committed to the number of opportunities for that error, i.e.,

$$\textbf{HEP} = \frac{\text{number of errors}}{\text{number of opportunities for error}}$$

There is, therefore, a close relationship between HR and human error concepts. The latter is used as a means to estimate the former (Leplat, 1985). Also, it is worth noting that HR is always evaluated with reference to a technological or man-machine system (MMS). The overall reliability is the product of HR and equipment reliability. Miller and Swain (1987), state that the purpose of HRA is to analyse the man-machine system and predict the potential for human error. This can be done qualitatively or quantitatively. Qualitative HRA usually involves the use of a task analysis to identify the possible errors and evaluate their consequences. In contrast, a quantitative HRA estimates the effects of human error in a system by assigning probabilities of failure and uncertainty bounds (UCBs) to the individual task elements and predict the overall probability of failure for a task. HRA can be used in probabilistic risk assessment (PRA) where the resulting probabilities of failure for task sequences are combined with probabilities of failure for the equipment in a system fault or event tree.

5.5 Human Error Definition

A necessary step to any human reliability/error analysis is the understanding of human errors nature.

Human errors have been defined according to the following points of view (Rasmussen, 1982, 1987; Rasmussen and Leplat, 1984):

(a) Human Errors: Causes of accidents

Errors can be described as causes of unfulfilled system purposes. If system performance is below the accepted standard, due to a human act or inact- the cause will very likely be identified as a human error. Rigby (1970), defines human error as any member of a set of human actions that exceeds some limit of acceptability, where the limit of acceptable or tolerable performance are defined by the system. In an objective assessment of human error, there is no connotation of blame or fault. If an error is made, it is necessary to find out *why*, i.e., to identify the underlying causes of the error so that the probability of recurrence of that error is reduced or eliminated. Few human errors, however, cause damage or lower the availability of systems because the potentially adverse effect of the errors is prevented or compensated for by other components or systems or by other human actions. These factors are known as *recovery factors* (Swain and Guttmann, 1983).

(b) Human Errors: Man-machine mismatch situations

For improvement of safety, a more fruitful point of view is to consider human errors as instances of man-machine/task misfits. It is more important to find the nature or dimensions of the misfits than to identify their causes. It is necessary to find *what* went wrong.

(c) Human errors: Experiments in an unkind environment

Basically, human errors should be seen as a result of human variability and versatility/adaptability. To optimise performance, it is important to have opportunities to 'cut corners', to perform trial and error experiments, and human errors can in a way be considered unsuccessful experiments with unacceptable consequences. Human acts are classified as human errors because they are performed in an 'unkind' environment where it is not possible for the human to correct the effects of inappropriate variations in performance before they lead to unacceptable consequences. Typically, because he either cannot immediately observe the effects of his 'errors', or because they are irreversible.

Human errors have also been defined as mistakes or slips.

Human Errors: 'Mistakes', 'Slips'

An attentional framework for the analysis of human errors has recently been provided by Reason (Reason and Mycielska, 1982; Reason, and Lucas, 1984) and Norman (Norman, 1981) in which they differentiate between errors of intention:*mistakes* and errors of action:*slips*. Human intentions and the resulting actions may be correct from the performer's point of view, from the goal he selects, but the effects on the system of his performance may be undesirable (mistakes). Most errors are, however, unintentional (action slips) -- the error just happens; it was not intended. A human error is simply an act which is counter-productive with respect to the performer's intentions or goals.

5.6 Human Performance Modelling

The importance of modelling the man-machine system (M-MS) has long been recognised, and many attempts have been made to develop human performance models for use in PRA (for a detailed review of the major models in this area, see, for example, Watson, 1985; Woods and Roth, 1986). What is a 'good' model depends on what the model is being used for (Pew and Baron, 1983). Bainbridge (1986), argues that models which describe and predict human behaviour may be much simpler than models which represent underlying mechanisms.

The main objective of modelling human performance for PRA is to provide data for design, training, and procedures decisions by either predicting or explaining human behaviour in normal and abnormal situations.

From the perspective of human reliability evaluation in PRA, modelling of human errors is of considerable importance. The first task of the HRA expert, regardeless of the HRA technique used, is the qualitative analysis of human errors, i.e., the modelling or identification of the potential errors within the particular situation considered. This has to be done with reference to a model of human performance and usually involves the use of task analysis and logical diagrams such as fault and event trees.

5.6.1 Task Analysis

Task analysis is a formal methodology, derived from systems analysis, which describes and analyses the performance demands made on the human elements of a system. Its goal is to provide the basis for integrating human and machine into a total M-MS (Drury et al., 1987).

There is no single method of task analysis applicable to all jobs or tasks. Instead, many methods and forms of task analysis exist with different approaches and objectives (see for example Singleton, 1974; Drury et al., 1987). All, however, have the common goal of identifying the performance demands of tasks. Task analysis is usually used in HRA to describe and analyse systematically the operator activities, the possible deviations (errors), and the underlying psychological functions.

One example of task analysis method recently developed is what has come to be known as hierarchical task analysis (HTA). HTA was originally developed by Annett and Duncan (1967) as an aid in training decisions. The technique was refined and demonstrated by Duncan (1974), Shepherd (1976), and Piso (1981). HTA has been found useful for application in the process industries (Embrey, 1981). HTA starts with the overall objectives of the system or task and describes an "operation" that fulfills these objectives. The operation may be broken down into sub-operations plus a "plan" that defines how these suboperations are linked. The plan in HTA is crucial since the difficulties facing the process operator may be completely overlooked if the analysts uses an approach which concentrates on "what" should be done, without systematically examining "when" things should be done. Each suboperation may in turn be broken down and described further. The stopping rule (P x C), proposed by Annett and Duncan is that further redescription is unnecessary where the product of probability of inadequate performance (P) and the costs of inadequate performance (C) is acceptable.

5.6.2 Human Performance Models

(i) Stimulus-Response Models - The Human as a System Component

The theoritical approach to human performance modelling frequently adopted by human factors specialists for the purpose of PRA is based on a stimulus-response model and is similar to the mechanistic and analytical one applied by system engineers,

i.e., by analogy with system components, the human has been treated as part of or a "component" in a closed-loop (input-output) system. This conventional treatement of the human in a M-MS began after World War II and was influenced by the behaviourist psychology approach to human performance modelling and the then new field of engineering psychology.

The role of the human in a M-MS has, in general, been described as follows (Swain and Guttmann, 1983):

1)- perceive information (via signals, signs, and environmental conditions) about the system state;

2)- process that information to determine what action, if any, is required;

3)- take the action decided upon.

As shown in Figure 5.1, the human is part of a closed loop system, since information about the effect of his outputs to the system is fed back (through displays and other external indicators) to become another human input. Thus, it is possible for the human to descriminate any significant difference between the actual system status and the desired system status if proper feedback is provided.

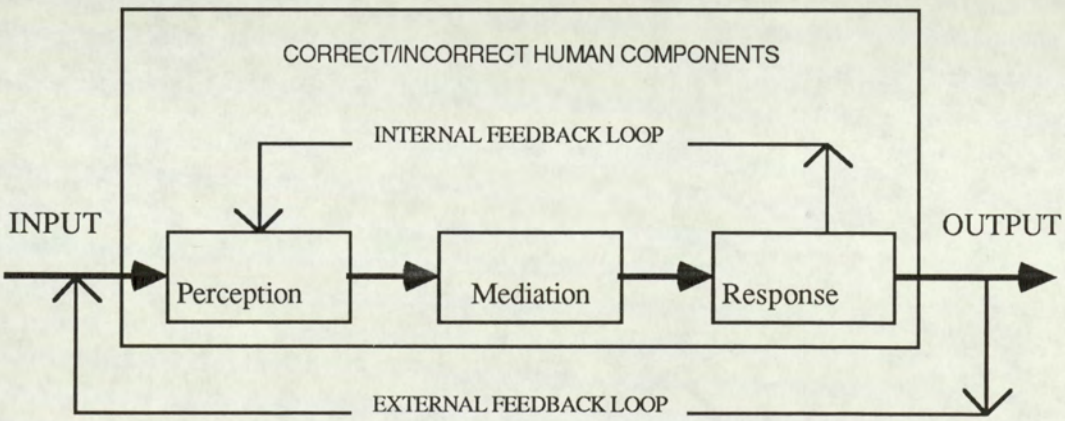


Figure 5.1 A Simplified Form of an Input-Output Response Model

(Adapted from Swain and Guttmann, 1983)

The model assumes that the human can be treated analytically, much as are other components in a system and concludes that the human behaves in the same manner as any other system component. However, considering the greater variability and interdependence in human performance, Swain and Guttmann (1983), comment that such a conclusion is invalid. They point out that:

> *"The human is an extremely interactive and adaptive component in any (MMS). Any attempt to depict his role in as simple a drawing as Figure (4.1) will introduce certain discontinuities in the representation of human behaviour that really do not exist."*

They, however, emphasise that:

> *"still, human failures can be studied objectively and quantified within limitations, as can any other component failure"*

and that:

> *"It is not necessary to understand fully the underlying dynamics of human behaviour in order to estimate the probabilities of human error"*

They, therefore, conclude that:

> *"for purposes of PRA only,* (this)*general descriptive model of the human will be sufficient."*

For the simple proceduralised situations, this model, which is only concerned with externally observable actions that have or could have an impact on the system or surrounding environment, is acceptable. However, human operators in nuclear and process plants perform a variety of mental or cognitive activities such as diagnosis and decision making, and there has been an increased realisation that these high level cognitive functions are of critical importance to nuclear and process safety and reliability, particularly in response to abnormal situations. These cognitive functions

are activated by subjective factors, such as intention, expectations and goals. They cannot be directly observed but must be inferred from characteristics of the task and the work situation together with the external manifestation of the error (Rasmussen, 1987) and, therefore, cannot be adequately captured or assessed by simple behaviourist input-output models.

There is, therefore, a need for models which include cognitive activities important to nuclear and process safety and reliability and that allow a more complete representation of the interactive functioning of the human as a system component.

(ii) Cognitive-oriented Models

To overcome the basic shortcomings of the engineering-biased models discussed above, various cognitive behaviour models have been proposed that can contribute to an understanding of cognitive activities in complex situations. Cognitive models attempt to describe, explain, and predict human behaviour by modelling what information people gather, how they acquire it, how it is represented internally, and how it is used to guide behaviour (Woods and Roth, 1986).

This section is concerned with the description and discussion of the major cognitive models that have or could be applied to human error modelling in nuclear and process industries.

A. *Rasmussen's Model:* - Cognitive Control & Human Error Mechanisms

Although a number of cognitive human performance models have been proposed which could be applied to human error modelling in process plant situations, (several comprehensive overviews and syntheses of literature in this area exist, see, for example, Sheridan et al., 1982; Woods and Roth, 1986), probably the only one which

has achieved widespread acceptance and use amongst HRA analysts has been the *three-level model* developed by Rasmussen (Rasmussen, 1980, 1982, 1987) and his co-workers at Riso National Laboratory in Denmark.

Rasmussen argues that humans are not simply deterministic input-output devices, but goal-oriented creatures who actively select their goals and seek the relevant information (Rasmussen, 1986). He distinguishes between:

(i) Three categories or levels of human behaviour/performance: *skill-*, *rule*, and *knowledge-based* behaviour. These levels and a simplified illustration of their interrelation are shown in Figure 5.2. The three levels in a real situation interact in a much more complex way than shown.

(ii) Three types of corresponding data/information percieved (Rasmussen, 1986):

a) *Signals*. Direct physical time-space data with no meaning.

b) *Signs*. Information that serves to activate or modify predetermined actions or prior experience.

c) *Symbols*. Whereas signs refer to percepts and rules of actions, symbols refer to concepts tied to functional properties and can be used for reasoning. They are defined by and refer to the internal, conceptual representation that is the basis for reasoning and planning.

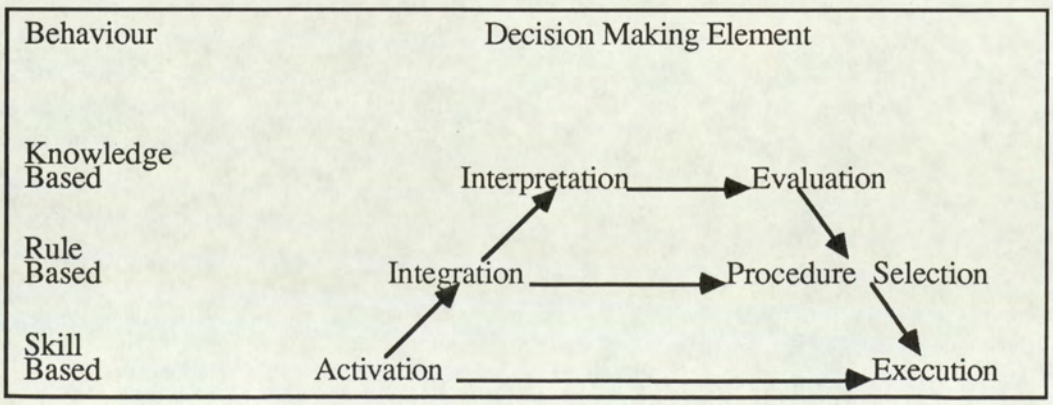


Figure 5.2 Decision Making Model (From Rasmussen, 1980)

The activation element corresponds to the initial alerting of the operator that action is required, e.g., by the onset of a pattern of symptoms or alarms. In most cases, with an experienced operator, these signals will be recognised as a familiar one, for which a simple pre-programmed action is available. Hence, the operator will normally directly execute the action appropriate to the symptoms. Behaviour of this type is called 'Skill Based', because it is characteristic of skilled, experienced operators.

If, as in some cases, the operator (or operating team) does not immediately recognise a pattern of indications, this will normally initiate a process of scanning and information collection. This will then usually lead to the selection of an appropriate procedure (either formalised as an explicit plant procedure, or from the operator memory) which will be executed to achieve the desired result. This mode of behaviour is referred to as a 'Rule Based', since it involves explicit rule-following.

If, even after the information gathering carried out during the integration phase, a diagnosis cannot be made which leads to the implementation of a specific procedure, the operator enters the realm of 'Knowledge Based' behaviour. In this mode, the decision maker cannot rely on past experience, but has to formulate new hypotheses from scratch. Once an interpretation has been arrived at, and a suitable strategy has been formulated to achieve a desired goal (e.g., a safe stable state) an appropriate procedure has to be selected to reach this goal. This may involve the use of existing procedure or the improvisation of a new one. Finally, the execution of the procedure will involve concrete actions such as opening and closing valve, turning on pump, etc.

The 'step ladder' form of the model arises from the fact that an operator will have only to go through all the stages in the model if he/she is faced with a totally unfamiliar situation or is inadequately trained or experienced. In most cases, 'short cuts' will be taken across the step ladder. These are normally appropriate, but can give rise to errors

in certain situations. The application of the model to operational errors in the nuclear industry is described in Rasmussen (1980).

Despite the widespread use of Rasmussen's model and its resultant taxonomy in nuclear and process industries, its validity is not universally accepted. Bainbridge, for example, believes that this model is not adequate as a model for human behaviour to be used in interface design (Bainbridge, 1984), but,

> *"it is useful in thinking about human error, and has proved particularly successful as an educational model, explaining cognitive processes to non-experts."* (Bainbridge, 1986).

Another criticism came from Hale and Glenden (1987), and was over the exact boundary between the three levels of behaviour. These authors suggest that:

> *"It is better to think of knowledge-based and skill-based behaviour as two ends of a spectrum, with rule-based behaviour as a somewhat ill-defined region half-way in between".*

They, however, conclude that:

> *"Despite this slight vagueness the model provides a valuable contribution to a systematic understanding of errors and incidents."*

Rasmussen (1987) himself comments that:

> *"...in present-day control rooms,...the context in which operators make decisions at the knowledge-based level is far too unstructured to allow the development of a model of their problem-solving process, and hence, to identify typical 'error' modes, except in very large terms..."*

He concludes that:

> *"Before (human) behaviour can be modeled, a systematic description of his decision-making context must be found (and), realistic models will probably only be possible, if his choice of goals and strategies is more constrained and controlled..."*

B. *Reason's Model:* - Generic Error-Modelling System (**GEMS**)

Reason makes a distinction between two broad classes of errors: slips and mistakes (these have been discussed in a previous section). Slips are departures of action from intention or errors which result from some failure in the excution stage of an action, whereas mistakes are deficiencies or failures in the judgemental and/or inferencial process involved in the selection of an objective or in the specification of the means to achieve it, i.e., planning failures (Reason, 1987a).

According to Reason (1987a), there are three psychological processes involved in planning:

a)- a working *database*, which corresponds to the working memory/ attention/consciousness constructions of information-processing models,

b)- a set of *mental operations* that act upon this database. These involve: selection, judgement, and decision making, and

c)- the *schemata*, or the structures comprising the long-term knowledge base, which control the database.

Reason (1987a), argues that from what is known of the heuristics influencing judgement, the limits of human information processing, and the generally agreed characteristics of schema function, it is possible to make a number of predictions about when distortions (or biases) will be introduced into the planning process, and what form they will take. He concludes that these sources of distortions are likely to lead to an inadequate database, unrealistic goal setting, faulty judgements of the likely consequences of actions, and an unwarranted confidence in the efficacy of the resulting plan. Figure 5.3, shows a general framework proposed by Reason (Reason, 1987b) for locating the principal limitations and biases giving rise to the more predictable types

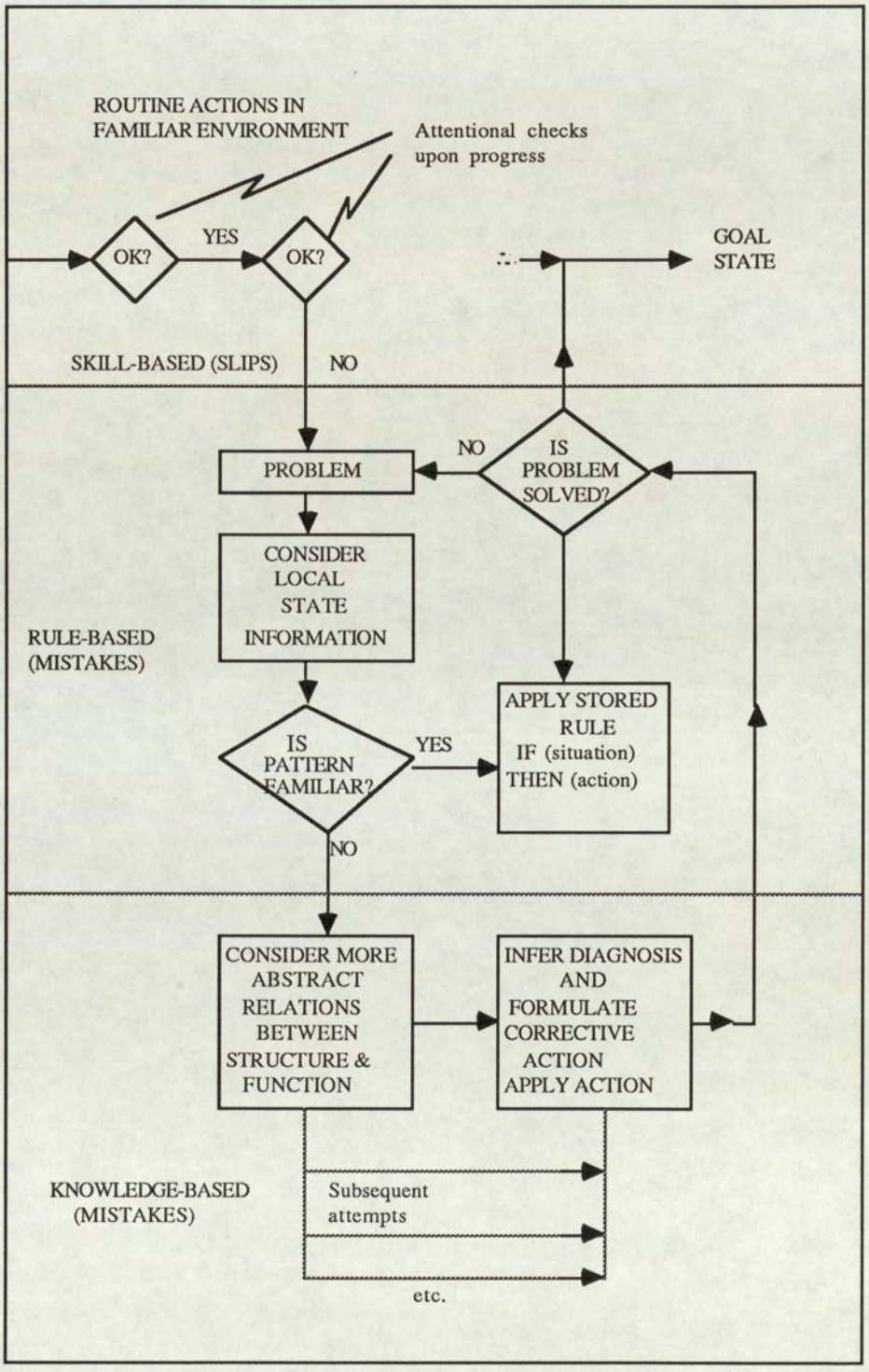


Figure 5.3 Dynamics of Generic Error-Modelling System (GEMS)

(from Reason, 1987b)

of human errors (skill-based slips, rule-based mistakes, and knowledge-based mistakes).

Although the basic structure of the two cognitive models discussed above is very similar, there is, however, no simple one-to-one mapping between these frameworks. Reason's distinctions relate to classical psychological concepts such as short- and long-term memory. Rasmussen's distinctions are derived from information processing and control concepts. The framework by Rasmussen is directed towards a clear separation of the representation of structural or declarative from procedural information. In Reason's approach this distinction is not sought since the focus is on psychological mechanisms (see Rasmussen et al., 1987).

5.6.3 Human Performance Shaping Factors (PSFs)

In modelling human performance, HRA analysts take account of those factors that have the most effect on performance. Many factors affect human performance. Swain and Guttmann (1983) differenciate between PSFs that are external to the person and include the entire work environment or task, and those PSFs that are internal and represent the individual characteristics of the person -- his skills, motivations, and expectations. The impact of internal PSFs on human reliability is generally reduced by an adequate training (Miller and Swain, 1987).

Stress is one of the most influencial internal PSFs. Stress is the body's physiological or psychological response to an external or internal stressor (Miller and Swain, 1987). Stress results from a work environment in which the demands placed on the operator by the system do not conform to his capabilities and limitations (Swain and Guttmann, 1983).

Watson (1985), considers that Error correction (recovery)/feedback is probably the most fundamental PSF and has implications at many levels and is consequently being incorporated much more explicitly in actual analytical models .

5.7 Human Errors Classification

In order to collect data, useful for understanding the way errors are caused and can be predicted and prevented, HRA experts categorise errors with reference to a taxonomy or classification scheme.

There are, however, many possible ways to describe and categorise human errors: it can be done, for example, with reference to task characteristics, to human psychological mechanisms, to environmental factors, etc. As a result many and varied taxonomies or classification schemes have been proposed. Some of these taxonomies were useful for psychological research but not very useful for HRA because they referred primarily to human variables. Other taxonomies brought in equipment variables and are more useful for HRA (for more detail, see Swain and Guttmann, 1983). More recent attempts to develop a taxonomy have been specifically directed at HRA as part of a PRA of nuclear or process operations (Rasmussen, 1981; Rasmussen et al., 1981; Topmiller et al., 1982, 1983; Comer et al., 1983; and Swain and Guttman, 1983). Most taxonomies are not empirically verified, formally tested against some criterion, or compared with other taxonomies.

5.7.1 Swain's Taxonomy

The Swains's THERP technique makes use of the Air Data Store (DS) categories (Munger et al., 1962) but has substantially expanded its taxonomy because it has expanded its data sources. The categories developed for the DS were primarily based on controls and displays parameters and the dimensions appropriate to these (e.g.,

lever length). Because of their focus on equipment components, the DS categories are quite molecular. Although Swain used DS values when appropriate, most of the HEPs in his Handbook (Swain and Guttmann, 1983) are more molar than those for the DS, .i.e., function and task- rather than component- oriented (e.g., initiate checking function, failure to use checklist properly), and they combine equipment and task characteristics. In THERP, only those human errors that constitute incorrect inputs to the system are considered. That is, performing a task incorrectly, failure to perform a required task, or failure to perform in time. These incorrect *human outputs* are categorised into two broad types of errors: *errors of omission* and *errors of commission*, with finer breakdown as shown in Table 5.1. Any of the above errors or any extraneous act is considered to be an error only when it reduces or has the potential for reducing system reliability, system safety, or the likelihood that some other system success criterion will be met.

5.7.2 Rasmussen's Taxonomy

Rasmussen (1987), notes that taxonomies generated in industrial systems are often aiming at a classification scheme suited to the analysis of existing human error data. This is to provide a basis for improvement of system reliability and for predicting human reliability during operation of new systems. Such an approach may succeed when the rate of change in technological development is slow and the element of operator tasks and interface designs therefore rather stable. In these circumstances, the taxonomy used by system designers may be rather simple, consisting of questionable types of 'task elements'. The model implicit in such taxonomies is a model of the task rather than the person. However, Rasmussen argues that the present rapid technological change in industry makes such simple taxonomies inadequate, and that human errors must be classified in terms of human characteristics:

Table 5.1 Categories of incorrect human outputs related to human reliability analysis (From Swain and Guttman, 1983)

ERRORS OF OMISSION

(1) Omits entire task
(2) Omits a step in a task

ERRORS OF COMMISSION

(1) Selection error:
    (a) Selects wrong contro l
    (b) Mispositions control (includes reversal errors, loose connections, etc.)
    (c) Issues wrong command or information (via voice or writing)
(2) Error of sequence
(3) Time error:
    (a) Too early
    (b) Too late
(4) Qualitative error:
    (a) Too little
    (b) Too much

> *"The emphasis must shift from task to the man-task mismatch as a basis for analysis and classification of human errors. The development of taxonomies must encompass the analysis, not only of manual task elements, but also of internal cognitive task components, and the psychological mechanisms associated with both. There is a severe need in systems design for data on error mode and probabilities during more complex decision tasks involved in emergencies."*

Rasmussen has developed what can be considered the most sophisticated taxonomy for use in the analysis of human errors in complex installations such as nuclear and process control plants. This taxonomy is based on the model described earlier (see Figure 5.2) and has been described in more detail in Rasmussen (1981), (1982), and Rasmussen et al.(1981).

The categories or dimensions of the taxonomy developed by Rasmussen are shown in Figure 5.4.

In the discussion of Rasmussen's taxonomy which follows, the categories are ordered according to the logical sequence of description and analysis of events involving human errors.

1- Personnel task

Identification of the task performed is important to characterise the circumstances during which the malfunction (error) occured. For detailed data collection purposes, the tasks involved must be analysed and the location of the failure in the task precisely identified. Furthermore, task analysis must be performed to determine the bias resulting from the potential for immediate error correction together with the frequency of error opportunities.

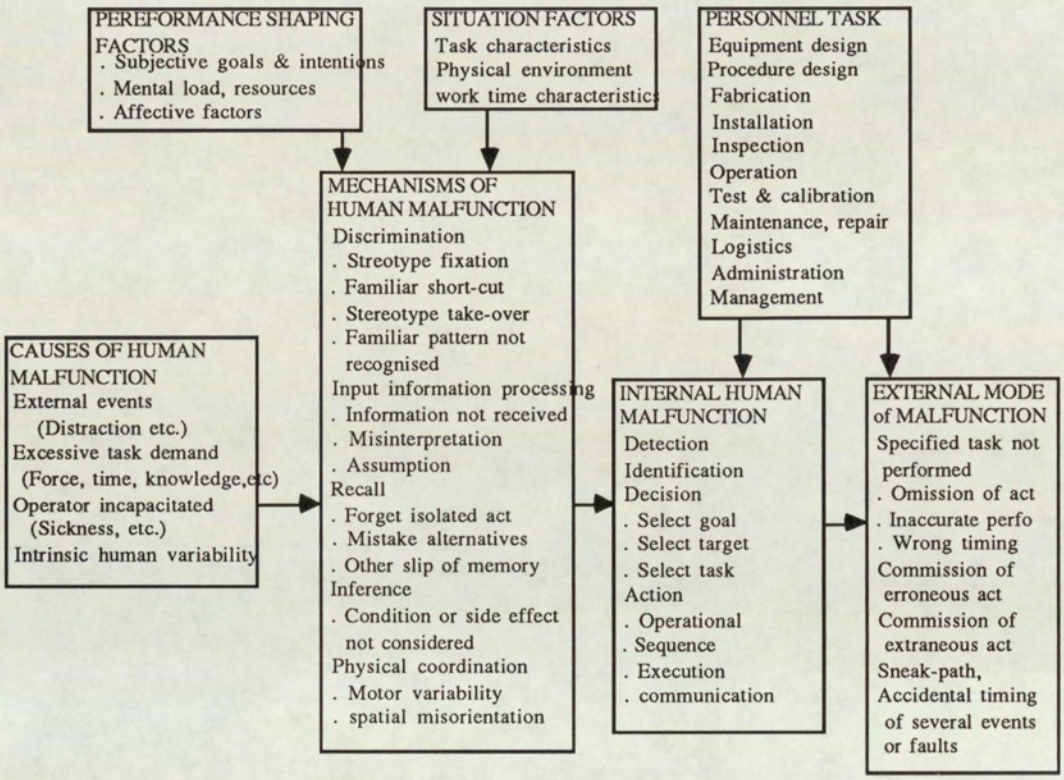


Figure 5.4 Multifacet taxonomy for description and analysis of events involving human malfunction. (From Rasmussen, 1982)

2- External mode of malfunction

This category describes the immediate, observable effect of human malfunction upon the task (or system) performance: this is the first element of a man-system mismatch. In this category, mismatch is expressed in terms of omission of acts in procedural sequences, acts on wrong components, reversals in a sequence, wrong timing, etc. Rasmussen (1987) notes that data from this category will be sufficient for design of systems applying technology very similar to the environment from which data are collected. When new technological means are introduced, however, data are needed which relate mismatches also to psychological mechanisms.

3 - Internal human malfunction - What went wrong?

To collect data which relate mismatches to psychological mechanisms, it is necessary first to characterise the mental task which has been involved in the event. A cognitive task analysis is therefore necessary to identify the internal mode of malfunction, i.e., the element of internal cognitive decision process which was affected, either by not being properly performed or by being bypassed by a habitual short-cut. Given the knowledge of the task and the particular external error mode found, it is in general possible *for the more familiar routine tasks* to judge the internal mode of malfunction from a case study (Rasmussen, 1980). For performance in unfamiliar situations, interviews and discussions with the performer are required. A systematic guide to the analysis of simple routine event reports, to identify *"what was wrong"* has been proposed (Rasmussen, 1982) and is shown in Figure 5.5.

4- Mechanisms of human malfunction - How it failed?

To further characterise an event in the category *mechanisms of human malfunction* with reference to psychological mechanisms, it is necessary to refer to the model of

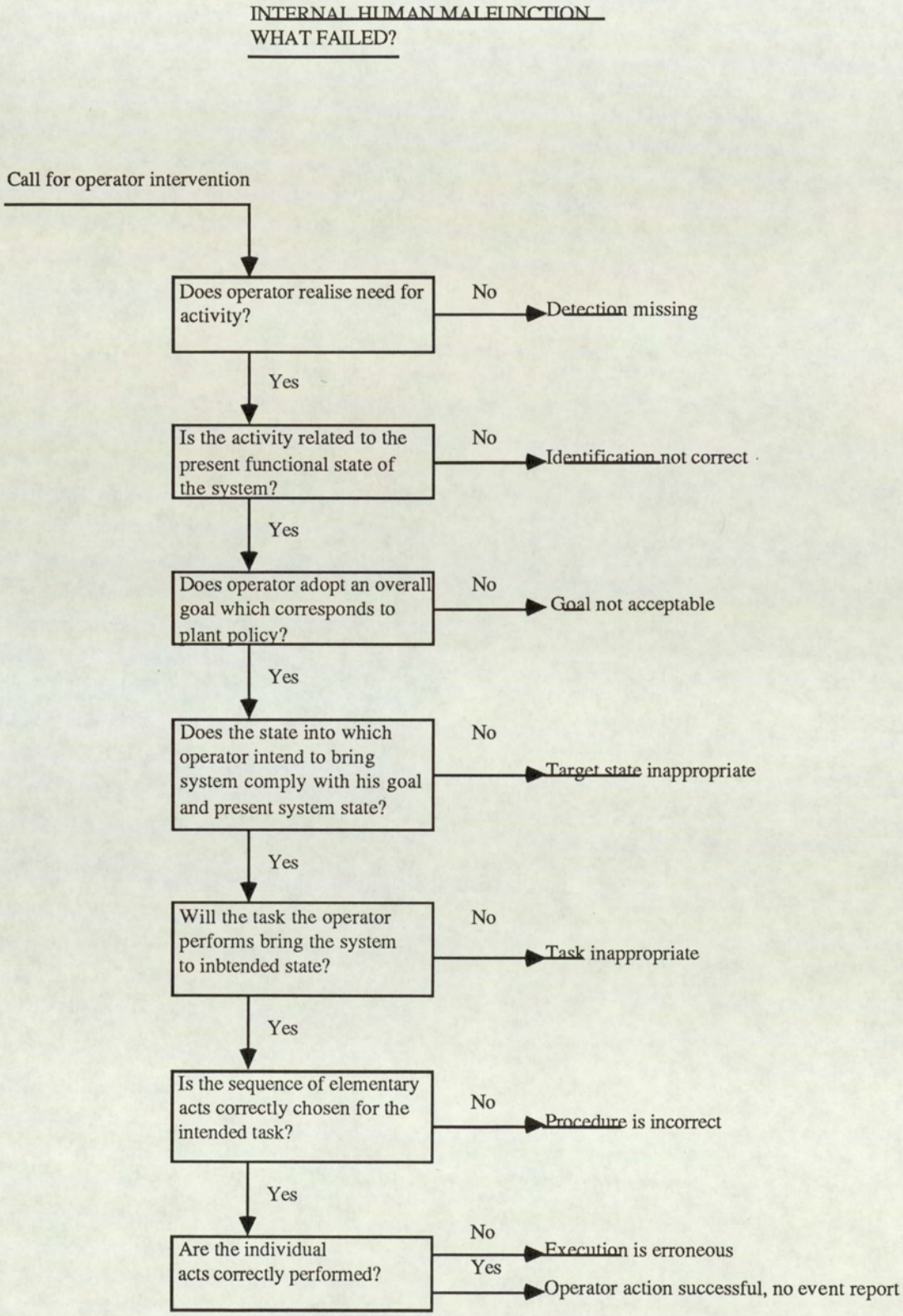


Figure 5.5 Guide to identify the internal human malfunction from event analysis.

cognitive control and human error mechanisms discussed in a previous section (see Figure 5.2), which can explain the mismatch characterised so far. The elements of this category are related to the 3 categories or levels of human behaviour: skill-, rule-, and knowledge-based, which are in their turn related to a decreasing familiarity with environment. Figure 5.6 shows example of guidelines derived from that model to be followed during the analysis (Rasmussen, 1982).

5- Causes of human malfunction- Why did it fail?

This category should identify the possible external causes of the inappropriate human action. As discussed previously, an error or malfunction implies a *change* from normal or expected function and this change can be due to a spontaneous internal human variability or a change in the external task conditions. Frequently, the human-system mismatch will not be due to spontaneous, inherent human variability, but to events in the environment, which act as precursors. Figure 5.7 illustrates a decision tree to guide data collection, proposed only as a framework ensuring consideration of major classes of causes.

6- Situation factors (events)

In the taxonomy, only *events* readily recognisable at distinct locations in time are considered causes, such as for instance interference by messages from colleagues, telephone calls, events in the environment such as bursts of noise, etc.More persistent conditions, like stress, bad design, inadequate instruction, etc., are considered performance affecting (shaping) factors and characterised separately.

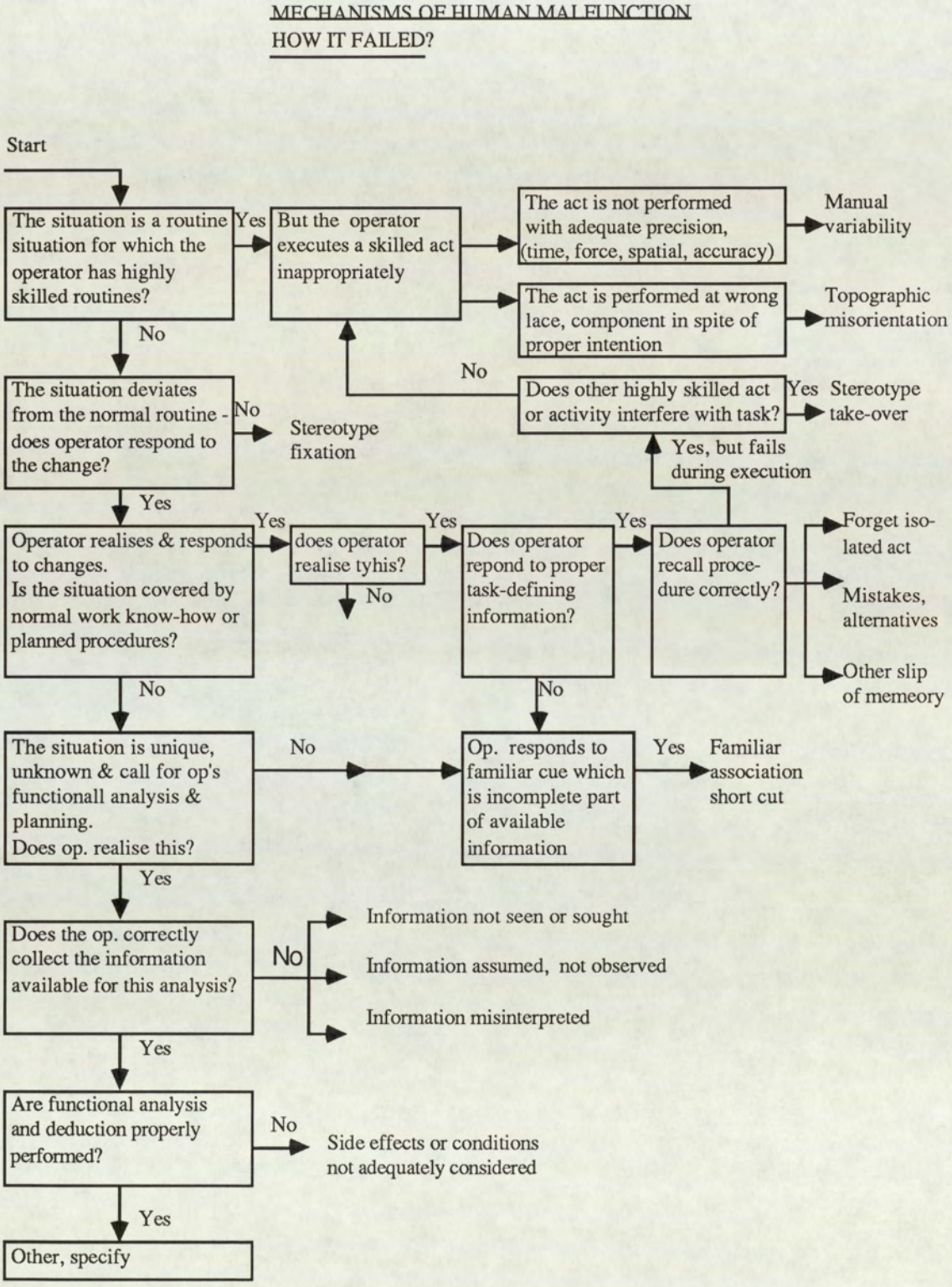


Figure 5.6 Guide for event analysis to identify the mechanism of human malfunction

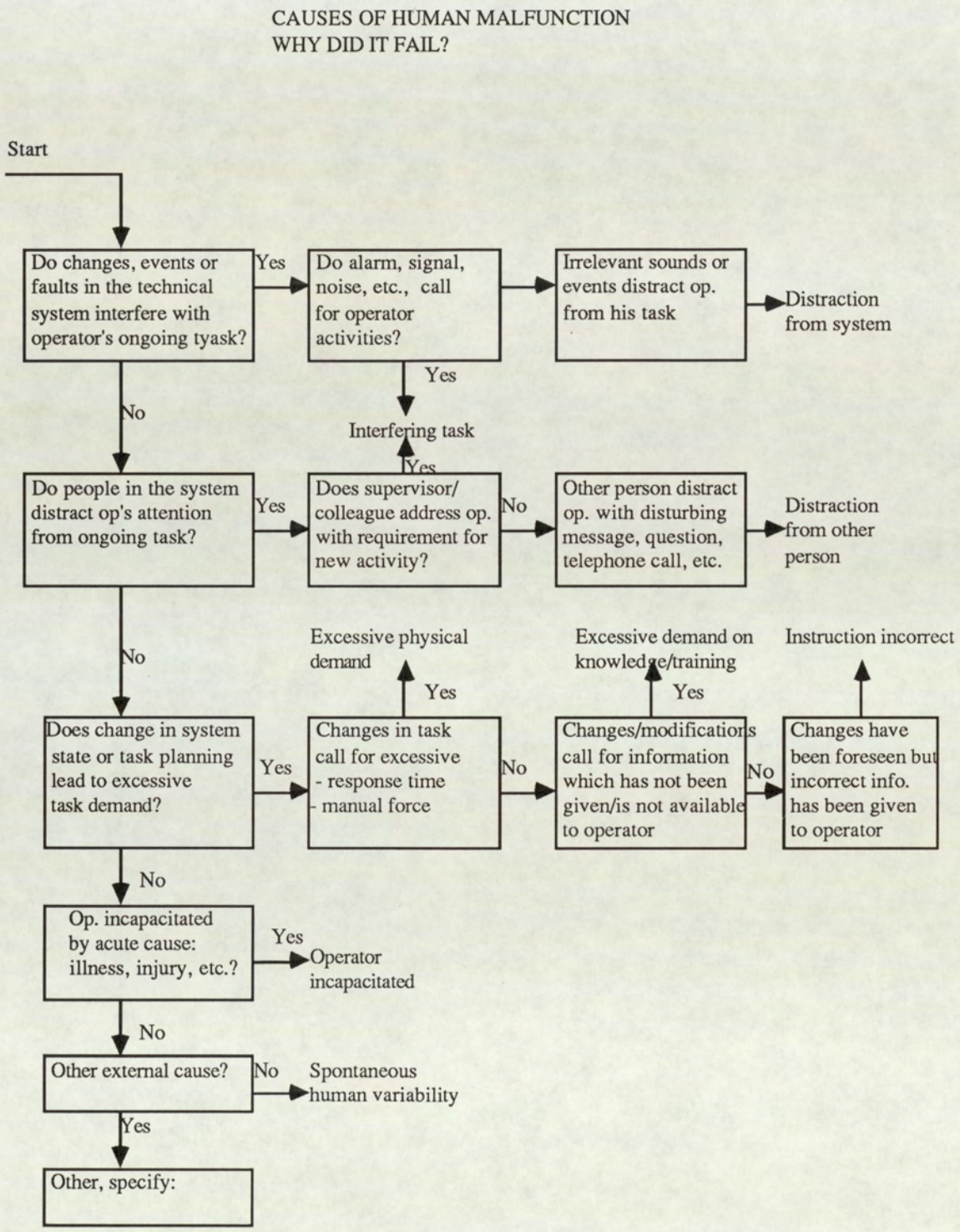


Figure 5.7 Guide for event analysis to identify external causes of human malfunction.

7- Performance shaping factors (PSFs)

This dimension of the taxonomy is intended to take into consideration that the human system interaction cannot be described adequately only considering the cognitive information processing level (see Figure 5.2). It includes general conditions which may influence error probablity, but do not cause errors. PSFs can only be identified by careful analysis.

This taxonomy, as emphasised by Rasmussen (1987), does not lead to a generic, hierarchical classification system, but to a multifacet system to characterise mismatch occurences. Some of its advantages are:

1. a very high resolution in the description of mismatch events can be obtained with only a manageable few elements in each of the dimensions of the taxonomy.
2. the internal structure of an event is preserved in the description, and can be regenerated in another context for analysis during design of new systems.

5.7.3 Embrey's Taxonomy

A recent variation of the Rasmussen's 3-category classification model is due to Embrey (Embrey, 1986, Embrey and Reason, 1986). His Decision Chart to identify task types is illustrated, slightly modified, in Figure 5.8. What Rasmussen called rule-based behaviour is split according to whether there is a diagnostic element, as in fault management during off-normal incidents. If diagnosis or decision making is needed but no rules are available to assist the activity, then operators must act based on deeper, more fundamental knowledge. Skills include pattern recognition and actions that are manual, well trained, well known, and practiced frequently. Otherwise, rule-based behaviour is elicited. This taxonomy has been extended by Reason (Reason, 1987) to

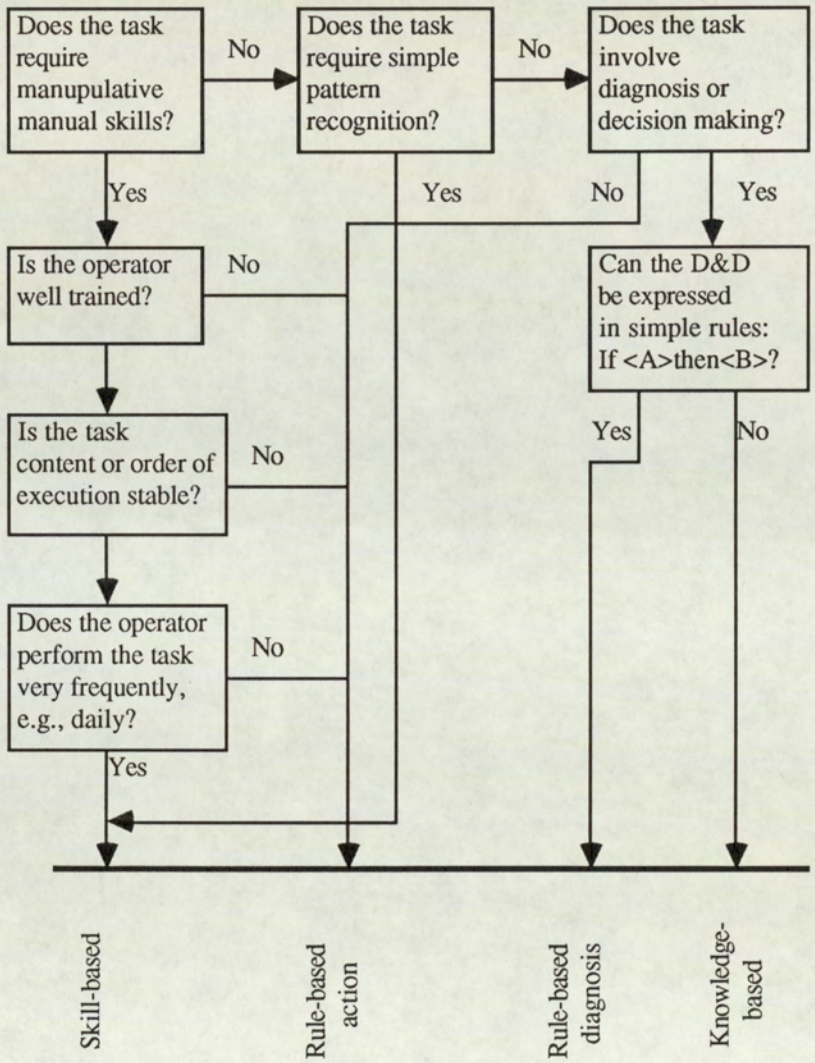


Figure 5.8 Embrey's taxonomy (from Dougherty & Fragola, 1988)

include a speculation as the role that intrinsic and extrinsic factors influence the types of errors that predominantly arise in the behaviour classes. Table 5.2 shows the relative contributions of intrinsic and extrinsic factors to skill-based slips, rule-based mistakes, and knowledge-based mistakes. Intrinsic factors include cognitive biases, attentional limitations, and limitation to human rationality. Extrinsic factors include task characteristics, effects of the situation, and factors from the environment.

Table 5.2 **Role of Intrinsic and Extrinsic Factors in Influencing Errors** (From Reason, 1987)

	INTRINSIC	EXTRINSIC
Skill-based Slips	high	moderate
Rule-based Mistakes	moderate	high
Knowledge-based Mistakes	low	very high

5.8 CONCLUSION

The main purpose of this chapter was to review the historical and theoritical basis of HRA, and identify and analyse the nature and characteristics of the HRA expert's tasks and expertise involved in performing the modelling or qualitative analysis part of the overall HRA procedure. The main conclusions that can be drawn from this analysis are as follows.

1. A necessary step to any human reliability analysis is the understanding of human error nature. This, however, is still ambigous. Although the nature of human error is still ambiguous, it can be considered as a deviation or an out-of-tolerance action with, unless recovered, unacceptable consequences on system performance. However, a more fruitful point of view is to consider human error as a result of

mismatch/misfit between man-machine/system or as an unsuccessful experiment in an unkind environment.

2. In order to provide engineers and managers with useful data for design, training, and procedures decisions, most HRA experts carry out analyses of the various tasks involved using methods such as task analysis, fault and event trees. These methods are usually based on models of human performance. The review of human performance models undertaken here, revealed two types of approaches used by HRA experts:

- First the conventional approach which is based on a stimulus-response model and that treats the human in an analogous fashion as hardware components.

- A second approach is mainly based on cognitive psychology theory as a means for understanding and analysing human performance. The models using this approach emphasise three major characteristics:

(i) They emphasise "knowing" rather than mere "response" as in the behaviourists' stimulus-response (S-R) models.
(ii) They emphasise mental structure/organisation.
(iii) The individual is viewed as being active, constructive and having intentions/goals rather than being a passive recipient of environmental stimulation.

3. When identifying or modelling human errors, it is necessary that the HRA analyst considers those factors (internal and external) that have the most effect on human performance (PSFs). These factors are numerous, but stress is considered to be one of the most influential PSFs.

4. Most of HRA experts recognise that there are different types of human errors with different consequences. They, therefore, describe and classify those errors that have been identified. Many approaches to human errors classification have been proposed and used by HRA experts, usually based on existing models of human performance. Swain's taxonomy is based on the stimulus-response or input-output model. The human ouputs have been categorised into two broad categories of errors: errors of omission and errors of commission. Rasmussen's taxonomy is based on his stimulus-cognition-response model and classifies errors according to three levels of behaviour: skill-, rule-, and knowledge-based behaviour.

The next chapter discusses the prblem of data in HRA, and provides a critical review of the conventional techniques developed and used for the quantitative analysis and prediction of human reliability. The purpose is to identify their state of development, to identify those that the proposed expert system will be based on, and to identify the HRA expert's tasks and subtasks necessary to perform them.

CHAPTER 6

CONVENTIONAL QUANTITATIVE APPROACHES TO HRA

6.1 Introduction

In the previous chapter, the historical and theoretical bases of HRA have been discussed, and the HRA expert's knowledge and expertise necessary to perform the modelling or qualitative analysis of human reliability have been identified. Some of the main questions addressed were: how HRA experts define human reliability and human error, and how they analyse, model and categorize human errors.

The main purpose of this chapter is similar to the previous one. That is, to define the nature of the domain problem. However, this chapter, first discusses the problem of HR data, it then attempts to answer, among others, the following important questions:

- what quantitative HRA techniques exist already?
- what is their state of development and complexity of use?
- what type of HRA method or methods should the proposed expert system program be based on?

In order to answer these question, a literature review of the existing human reliability prediction methods was performed. The analysis of the literature review that was accomplished was done in two steps.

Initially, a general literature review was performed which identified major types of HRA methodologies and techniques.

The methodologies that were not eliminated after step one were investigated further in step two and their description is presented in Chapter 7.

6.2 Human Reliability Data Problem

The most serious problem in the field of HR is the lack of data (Miller and Swain, 1987). However, the problem of securing data to serve as an HR data base is central to HRA. Meister (1982) states that:

> *"The data bank is a prerequisite if one is going to predict quantitatively and specifically".*

This section, first discusses the inherent problems of human error data collection and reviews the methods currently used. It then reviews the formal or published databases. Finally, the current efforts to develop future data bases are discussed.

6.2.1 HR Data Collection

One of the major problems associated with data collection is that of generalizability. In order for error data to be generalized, HRA experts make two assumptions: (1) Although every error made by a person is unique, HRA experts assume that there are enough similarities in error characteristics that it might be possible to combine similar errors into groups (see "Error Classification" in Chapter 5). (2) They also assume that errors that are similar from a behavioural and situational perspective have similar probabilities of occurrence.

A prerequisite for a data base is a task taxonomy, however, as discussed in Chapter 5, different taxonomies have been developed and there is no agreement among experts on which taxonomy is best and be used.

Another problem with human error data is the variation in the type of methods or sources from which data are collected. There are a number of ways in which human error data can be collected. The major ones are *manual* collection from the operational environment (the "real world"), *automatic* data collection in the real world, *self report*, *experimental studies*, and *expert opinion*. Each method has been utilized at some time or other.

i) Manual Collection

Presumes a human data recorder physically present in a facility who observes task performance and notes specified events (errors).

This method is expensive because it requires that at least one person be physically present at all times and a large number of data collectors in some situations. It is fallible because the observer may fail to recognize errors. He may not be sufficiently alert, the

event may occur too rapidly, or the error may be covert. In addition, his presence may create doubts into the workers minds concerning the reason for him being here, resulting in their deliberate concealment of the event and the error.

ii) Automatic Collection

Development of hardware/software (simulator) systems that automatically record events or actions so that the defects of manual method are avoided.

These systems are expensive to develop, subject to malfunction and most serious of all, limited to control actions so that the operator's perceptual and cognitive functions must be ignored unless they can be inferred from the control data.

iii) Self Report

In this situation, the one who made the error reports the error. Self Report can occur in various ways. For example, the operator may fill out a report form at the time the error is made, or he can report verbally to a supervisor who completes the form. The operator can be interviewed following the shift. A variation of the self-report method is the questionnaire survey.

There are tremendous difficulties with this methodology because workers are reluctant to confess making an error. Moreover, the individual may genuinely forget to report or, if he does report, he may not have noted all the circumstances about the error that are desired.

iv) Experimental Studies

There are two types of such data sources: (1) those performed specifically by the researcher to study human reliability and (2) the general psychological literature.

The largest amount of data available comes from laboratory experiments. Unfortunately, laboratory studies are often highly unrealisic.

v) Expert Judgement

This data source is commonly used by all HR methodologists. Such judgements can be formal or informal. The formal ones employs psychometric techniques. The informal ones are purely judgemental and make no use of psychometrics. Meister (1985) states that it is probably not feasible to ask even admitted experts to make direct judgements of error probability, but one can ask judges to make paired comparisons or to rank tasks in terms of error probability.

The main advantage of formal judgements is that experts are usually available when all data sources are unavailable.

The disadvantages are obvious: 1) often experts are not as expert as they should be, 2) data based on expert judgement are likely to be less reliable than data gathered by other means, and their validity is suspect until verified by empirical data.

Nevertheless, all HRA developers have used expert opinion at one time or another. Expert judgements can be secured formally and systematically from judges but the most common use of expert opinion is when HR researcher extrapolates data or quantifies his hunches. The informal use of judgement is much more common than the formal.

Meister (1984) notes that:

> *"The HR researcher is, if nothing else, a pragmatist and makes use of all the methods derived above when necessary and/or convenient."*

6.2.2 HR Databases Review

Despite that the need for a quantitative data bank (DB) on human performance has long been recognized, quite few formal databases have been developed and these are very limited.

i) AIR Data Store (DS)

The primary DB is the DS (Munger et al.,1962). Over the years there have been a few attempts to add to the DS in an attempt to apply it to particular systems (for example, Irwin et al., 1964 and Meister, 1967). These databanks have never been used. The

Irwin et al.databank can be applied only to situations sufficiently like the original ones. Meister's databank is too restricted (Meister, 1985).

ii) Technique for Human Error Rate Prediction (THERP)

The THERP method developed by Swain at the Sandia National Laboratories (Swain, 1964) makes use of the DS categories but has substantially expanded its taxonomy because it has expanded its data sources. The "Handbook" recently published by Swain and Guttmann (Swain and Guttmann, 1983) provides a great number of human error probabilities (HEPs). The HEPs were derived from actual errors in nuclear power plants, training simulators, job situations in process control industries and military situations; also experiments and field studies with real world tasks; data from the psychological literature; and expert judges.

The DB developed by Swain is the only one that has been expanded and modified over the years; all the others have been frozen as they were published. With regard to these error probabilities, the"Handbook", presents the best data available (Heslinga, 1983).

6.2.3 Current Studies

With the increased interest in HR prompted by the TMI nuclear incident, several attempts are being made to compile and analyse all available data banks.

One study supported by the Nuclear Regulatory Commission (NRC) was to summarize and publish all available DBs (Topmiller et al., 1983), to find any additional data sources and to develop a methodology for human error data collection at nuclear power stations. The effort to find additional data sources has been somewhat disappointing. The Institute of Electrical and Electronic Engineers (IEEE), (1981) made a survey of models and data bases relating to human performance in aerospace,

military, fossil fuel and nuclear power sources and concluded that *"although it may seem...that a significant amount of data exists, this is far from the case"* (IEEE, 1981).

A second study sponsored by the NRC has attempted to validate the HEP values in Swain and Guttmann (1980). Beare and Dorris (1983) collected error data from trainees in nuclear control room simulators and compared those with the Swain/Guttmann HEPs observed error rates for errors of commission were in close agreement, but errors of omission were not. Meister (1985) comments that the latter discrepancy may have resulted from the artificial nature of the training exercises.

For more detail, see Meister (1984), (1985); Miller and Swain (1987).

6.3 Conventional HRA Approaches

The general review of HRA methodologies that was performed identified the following three major categories of approaches:

(a) "Analysis and Synthesis", "Reductionist" or "Decomposition" approaches.
(b) "Time Dependent" approaches.
(c) "Subjective Expert Judgement" approaches.

These general categories are briefly described and discussed below (for more detail see Embrey, 1976; Meister, 1984, 1985; Miller and Swain, 1987).

The first category depends heavily on task analysis which involves the breaking down of human tasks into individual behaviour units, such as operating switches, closing valves or reading instruments. A probability of error is then assigned to each behaviour element. The probability of error for more complex tasks is then calculated by the combination of basic elements in a sequence of sub-tasks. The approach most

prominent in recent years is THERP (Technique for Human Error Rate Prediction) (Swain et al.,1983). These approaches are most suitable for simple procedural tasks, their application to highly dynamic, decision-making tasks has, however, been questioned.

The second category attempts to apply classical theory of reliability time dependent modelling, to predict human performance. Much of the early work on modelling HR as a function of time was carried out by Askren and Regulinski (e.g., Regulinski and Askren, 1969, 1972, Askren and Regulinski, 1969a, b, Regulinski, 1973) who derived a general mathematical model of human performance reliability that closely conforms to classical reliability methods and therefore could easily be combined with equipment reliability.

Most of the new time-dependent methods use response-time correlations or time reliability curves (TRC) to address the quantification of human contributions, especially those associated with detection, diagnosis and decision-making. Amongst those are the Operator Action Tree (OAT) (Hall et al, 1982), the TRC contained in the latest version of NUREG/CR-1278, (Swain et al.,1983), and the Human Cognitive Reliability (HCR) correlation (Hannaman et al., 1984). The main principle is, given zero time to respond, the HEP necessarily is unity; given long times, say 30 to 40 h, a HEP should be minimal. But Edwards, (1984), notes that:

> *"humans perceive a time different from clock-time; and if an action is based on a faulty diagnosis, then added time may actually lead to an increasing HEP."*

The third category makes a much greater use of quantified expert judgement, supplementing the currently inadequate human error probability data base for various types of tasks. The main characteristic of these techniques is that the task is not decomposed into smaller units to which HEPs are applied from a data bank. Instead,

the task is evaluated as a whole and quantified by either making absolute or structured judgements of the required HEP. The most developed of these approaches is the Success Likelihood Index Methodology (SLIM) which is described in Embrey et al., 1984). A review of subjectively based techniques for HRA is provided in Stillwell et al. (1982) and Seaver and Stillwell (1983).

Table 6.1 lists some of the major HRA methods as well as a list of their authors or reference sources.

Table 6.1 Major Conventional Human Reliability Analysis Techniques

Technique	Author(s)/Source(s)
APJ: Absolute Reliability Judgement	(Comer et al., 1984)
HCR: Human Cognitive Reliability	(Hannaman et al., 1984a)
HEART:Human Error Assessment & Reduction Technique	(Williams, 1986)
MATHEMATICAL Model	(Askren & Regulinski, 1969)
OAT: Operator Action Tree	(Hall et al., 1982)
PC: Paired Comparison	(Blanchard et al., 1966)
SHARP: Systematic Human Action Reliability Procedure	(Hannaman et al., 1984b)
SIMULATION Methods	(Siegel et al., 1974)
SLIM: Success Likelihood Index Method	(Embrey et al., 1984)
STAHR:-	
Sociotechnical Approach to Assessing Human Reliability	(Phillips et al., 1985a,b)
TESEO: Technica Empirica Stima Errori Operatori	(Bello et al., 1980)
THERP: Technique for Human Error Rate Prediction	(Swain et al., 1983)

6.4 Incorporation of HRA into PRA

As discussed in Chapter 3, HRA is usually carried out to support a larger-scale PRA study. Miller and Swain (1987) pointed out that:

> *"The process of integrating human-system interactions into the PRA framework should be systematic and preferably standardized across an industry."*

Many studies have been carried out to investigate this process. These studies are discussed below.

One major study was the project sponsored by the Electric Power Research Institute (EPRI), called *Systematic Human Action Reliability Procedure* (SHARP) (Hannaman et al., 1984b). SHARP's objectives were to provide a structured approach to the incorporation of human interactions into PRAs, and to enhance the documentation reproducibility and of the study.There was no intention to develop new HRA models or techniques, or to rely on any existing HRA approach.

SHARP is not considered a "cookbook," but a menu of steps, activities, and rules that can be selectively applied by the analyst in a PRA study.

Another document that discusses the integration of HRA into PRA is the *Interim Reliability Evaluation Program Procedures Guides* (Carlson et al., 1983). Although the report reviews analysis procedures predominantly associated with plant systems reliability and accident sequences, a major section deals with human reliability and procedural analysis methods and the role they play in the PRA.

Other major contributions in this area include the *Probabilistic Safety Analysis Procedures Guide* (Papazoglou et al., 1984), the IEEE/ANS *PRA Procedures Guide*

(NUREG/CR-2300, 1983), and *A Procedure for conducting a Human Reliability Analysis for Nuclear Power Plants* (Bell and Swain, 1983).

6.5 HRA Techniques Selection Criteria

Before examining the individual HRA techniques that were selected for implementation into the proposed expert system, the major criteria that have been proposed by HRA experts as a basis for judging the quality of existing HRA predictive techniques are presented.

Although various criteria for evaluating HRA techniques have been proposed, there is no general consensus among HRA practitioners on which set of criteria are best and be used.

The following set of criteria or desirable qualities for assessing any HR predictive technique, have been suggested by Meister (for more detail, see Meister, 1971):

(a) The technique could be usable by non-specialists.

(b) Tedious calculations should not be necessary.

(c) It should not require the application of performance data not readily available. It should not be necessary to perform experiments nor invoke expert judgement.

(d) It should be usable at all stages of system development and should handle all system elements in both molar (i.e., task orientated) or molecular (at the level of individual behaviours such as pressing a button) forms.

(e) Usable design recommendations should be produced.

(f) The reliability metric output by the technique should be understandable in terms of concrete operations.

(g) The technique must be validated by the collection of performance data in an operational setting.

(h) The predictive output of the technique must be compatible with those of equipment prediction techniques.

(i) The technique must be capable of assimilating data from a wide variety of sources.

The above list of qualities, although not an exhaustive one, have been considered by many HRA researchers to be important and/or desirable within a quantitative predictive technique (Embrey, 1976).

Two more important criteria have been identified by Siegel et al., (1983). The predictive technique should:

(a) include capabilities for handling cognitive task elements.

(b) provide means for sensitivity analysis of parameters within the methodology.

6.6 Discussion of Conventional HRA Methods

6.6.1 Assumptions and Complexity

The critical review of conventional HR quantification techniques undertaken here, shows that HRA experts have developed, used, and continue to refine a number of alternative approaches to the problems of human reliability modelling and quantification. However, none of these techniques in their present form constitutes an ideal methodology from the point of view of the desirable criteria outlined above.

In this section, the complexity of the conventional HRA methods and the amount of judgement and experience needed to perform them is briefly discussed by examining the main concepts used and assumptions made by HRA experts in solving problems in this field (for more detail, see Meister, 1984 and 1985).

Although the majority of the techniques used in human reliability work have been developed in the context of the nuclear industries, they can and have also been applied in a wide variety of other process control industries.

A major assumption made by most of the HRA analysts is that the human can be treated analytically like other system components.

All HRA experts assume that human performance is modified/influenced by PSFs. The list of such modifiers is very long and varies from environmental conditions to idiosyncratic characteristics of the human to attitudes and social/organizational conditions. The assumption that PSFs influence HR predictions requires the analyst to attempt to factor these into the predictions, but how adequately this can be done is not clear. Of the multitude of PSFs presumably influencing human performance, HRA analysts attempt to account for only a few, such as stress, proficiency and experience. The relative inability to account for other PSFs, such as motivation, reflects a lack of data about how PSFs affect job performance.

All HRA analysts explicitly or implicitly accept the notion that behaviour can be described in term of the stimulus-organism/cognition-response paradigm, which separates behaviour into input-internal processing-response segments. This is the conceptual basis on which most of the HRA experts use task analysis to break down behaviours at a molar level into their individual elements and then resynthesise them. During this decomposition process, it is necessary to decide which behavioural parameters are relevant to system objectives.

All the analysts (except the authors of the AIR Data Store) assume the interdependence of task parameters and tasks. Swain distinguishes among three types of probability: the probability inherent in the task when it is considered as an isolated entity, conditional

human error probability which is the error probability of a task given success or failure on other tasks, and the combination of these two probabilities. He further subdivides conditional error probability into five levels of dependency representing degrees of effect of one task performance on another.

Skill level requires another set of assumptions. Swain suggests that novices under optimum stress performing nonroutine tasks are twice as likely to make errors as experienced personnel (Swain and Guttmann,1983). The concept that less experienced personnel would have a higher HEP than more experienced personnel is *"entirely reasonable"* (Meister, 1985). However, there are few experimental data that support this premise.

A common assumption made by most HRA analysts is that there are different kinds of errors with different kinds of consequences. The most frequent types of error one would find in nuclear and process control systems are errors of detection, diagnosis and decision making.

A major assumption made by Swain is that human error, unless recovered or trivial, is equivalent to failure to accomplish a task or some system required activity. As Swain puts it, the basic measure of human performance is the HEP, which is the probability that when a task is performed an error will occur. *"The probability of successful performance of a task is generally expressed as 1- HEP"* (Swain and Guttmann, 1983). However, many errors do not result in task failure; they are either too insignificant or they are recovered.

The use of probability mathematics is common to all the HR methodologies and *"represents the recognition of the probabilistic nature of human performance"* (Meister, 1985).

6.6.2 Limitations of Conventional HRA Methods

The following are some of the general limitations and criticisms that have been identified and directed to current HRA techniques (see, for example, Meister, 1984, 1985; the US.NUREG-1050, 1984).

- Human behaviour is a complex subject that does not lend itself to simple models like those for component and system reliability. This makes the analysis of human behaviours more dependent on judgement of the analysts.

- An adequate data base does not exist with which to make accurate assessments and predictions.

- An unacceptably large subjective element is present in some HRA methods, making their predictions invalid.

- Another objection is that lacking empirical data, the analyst who attempts to apply THERP must be highly skilled in performing the necessary analyses before his quantitative estimates can be accepted. For example, in assigning HEP to the events depicted in the fault or event tree.

- Generic HEPs have been applied on a judgemental basis, because a simplified and comprehensive model including the various factors that affect human performance has yet to be fully developed.

- Human impacts have been described as binary success and failure states to match the logic used for equipment failures. This approach does not account for the full range of human interactions, such as time dependent effects.

- Considerable variability was observed between the studies (and to some degree within a given study) as to the selection and application of an approach to quantify the human error probabilities. Additionally, there were differences in interpretation of the data in the Handbook (Swain and Guttmann, 1983).

- One major finding of a more recent "critical review of analytical techniques for risk studies in nuclear power stations" (Kroger et al., 1987), is that:

"There is still a lack of a standardized, structured technique to derive data for individual actions from operating experience or estimates by experts."

6.7 The Selected HRA Techniques

The general literature review revealed that current state of development of HRA techniques do have limitations and difficulties in their applications. In the review of the individual predictive techniques, three major HRA techniques were selected which are used in the human factors (ergonomics) community to identify and quantify existing or potential human performance problems in man-machine interface. The methods selected for the proposed expert system are those that have been directly applied to HR problems in nuclear and process industries or are potentially applicable to them.
The three selected HRA methodologies are:

1)- Technique for Human Error Rate Prediction (THERP), (Swain and Guttmann, 1983).

2)- Success Likelihood Index Method (SLIM), (Embrey et al., 1984).

3) Absolute Probabilities Judgement (APJ), (Comer et al., 1984).

These techniques have been selected for the following reasons:

i)- THERP, because it is regarded as the most powerful and systematic methodology for the quantification of human reliability, is well documented, is the HR technique most often employed, as well as providing a human reliability and Performance Shaping Factors (PSFs) data banks.

ii)- SLIM, because it is the most developed subjective technique which illustrates how it is possible to systematically derive HR probabilities from expert opinion given two known probabilities of failure, and also because it does not require extensive decomposition of the tasks assessed.

iii)- APJ, because it is the most direct approach to the quantification of human error probabilities, which, unlike SLIM method, does not require the availability of two reference tasks (not always obtainable), and since the data provided by THERP may be not relevant, or applicable to all circumstances, the APJ approach can be used to estimate those missing HEPs.

These three techniques are described in more detail in Chapter 7.

6.8 CONCLUSION

Human reliability data remains a major problem in conducting PRAs. Despite the shortcomings of the THERP/Handbook, most of the analysts still relie on it as a guide to HRA, as well as a primary source of data on human reliability probabilities. However, many research programs are underway to gather data from other sources such as simulator exercises. In addition to gathering data to validate and supplement the data in the handbook, plans are underway to expand treatment of cognitive aspects of behaviour in process plants.

Meister (1985), comments that it is impossible to counter the objection that the data are too scanty and subjectivity too rampant.

> *"How much data would be enough? and should one wait until there are enough data? Because the HRA methodology is designed to respond to problems of the here and now, it cannot hibernate until some hypothetical time when there will be enough data."*

Although there may be a substantial discrepancy in any HR prediction made, the amount of error in that discrepancy is considered not to be as important as it would otherwise be. Even if errors cannot be precisely predicted, some idea of where they can be expected is needed. To reject the HRA method because of its deficiencies would be to become essentially impotent to predict human performance (Meister, 1985).

HR prediction is still a research and development activity. However, the last twenty years or so have resulted in HRA evolving into a credible and useful discipline that can be applied to risk analysis in nuclear and process industries (Dougherty and Fragola, 1988).

As discussed above, numerious HRA methods have been developed and used in nuclear and process industries - this is a fact. It is also clear that, despite the many steps that have been taken (particularly by EPRI and the USNRC) since the TMI accident to overcome their deficiencies, such as organization of the HRA process, and the development of acceptable models of operator cognitive processing, existing HRA techniques are still not fully developed.

Whereas the development of new techniques and the usage of existing HRA techniques in PRA for nuclear and process industries are increasing, a major obstacle to their broader usage has been the large level of effort required to select, amongst the various techniques available, and use the most appropriate technique to the situation of interest.

Therefore, tools are needed to assist the non-expert, in the selection and detailed application of the appropriate HRA to the specific situation considered.

The next chapter describes the three HRA techniques that have been selected for implementation in the expert system developed here to overcome some of the shortcomings of the conventional approaches.

CHAPTER 7

SELECTED HRA METHODS

7.1 Introduction

The main purpose of this chapter is to describe the three selected HRA techniques and more precisely to answer the following two questions:

1- what are the main human reliability expert's tasks, subtasks, and expertise that are essential in the performance of THERP, SLIM, and APJ?

2- what are the main advantages and disadvantages of the three selected HRA techniques?

Obviously, it is impossible to describe each technique in general detail here; for more information on these techniques and review of other HRA methods, see Meister (1971, 1984, 1985); Embrey (1976); Pew et al. (1977); and UKAEA, (1988).

7.2 THERP Method

7.2.1 THERP Background

As discussed in Chapters 5 and 6, human reliability analysis (HRA) evolved into a comprehensive discipline in the 1960's. The dominant approach to HRA at that time became known as the *Technique for Human Error Rate Prediction* (THERP).

THERP is an analytical technique currently known as *"Human Reliability Analysis"* (HRA). It is the oldest (developed in 1961 at Sandia Laboratories) and most widely used HRA method to date. A handbook which describes its use has been published,

first in a draft form (Swain et al.,1980) and then was revised and published (Swain et al., 1983) as a result of a peer-review by Brune et al., (1983). It has been described extensively by its primary developer (Swain) and his co-workers in a long list of reports, handbooks and lectures (see for example, Swain and Guttmann, 1983, Bell and Swain, 1983). THERP was first used in the Reactor Safety Study (WASH-1400, 1975) and is or has been applied to the analysis of human reliability in several subsequent PRAs in many agencies and countries (Dougherty, 1983; Kolb et al., 1982; Carlson et al., 1983).

Most of the applications of THERP have involved estimates of the probabilities that system-required tasks will be executed correctly and, in some cases, within specified time limits. Nearly all of the applications have assessed the probability that individuals or teams would carry out specified procedures correctly under varying degrees of stress. The types of tasks include assembly of equipment, air crews performing military tasks, and the various NPP activities assessed in WASH-1400 and subsequent PRAs of NPP operations. It is only with the most recent PRAs that the so-called *cognitive* aspects of behaviour have been addressed.

Since the use of THERP in WASH-1400, there have been some refinements in this method and the data it uses (Swain and Guttmann, 1983):

1. The development of several human performance models:

   (a) a model of positive dependence.
   (b) several models for the perceptual and response aspects of human activities.
   (c) an interim model to assess cumulative probability over time of correct diagnosis of transients and other abnormal events (i.e, the "cognitive" aspect).

(d) expansion of the extremely high stress model to include several other stress levels.

(e) a model for assessing team interaction in control rooms.

2. A method of estimating and propagating uncertainty bounds in an HRA.

3. The Handbook itself, which brings together the latest developments.

A shortened version of the THERP/Handbook approach to human reliability analysis for PRA has been developed by Swain under financial support from the US.NRC. This shortened version was prepared and tried out as part of the Accident Sequence Evaluation Program (ASEP), (Swain, 1987).

7.2.2 THERP Description

The acronym THERP was first used to designate the human reliability method developed at Sandia National Laboratories (Swain, 1963). Despite that the term *"Human Error Rate"* (HER) has been dropped in favour of *"Human Error Probability"* (HEP), the acronym THERP has been retained because it is now well established.

THERP is currently defined as a method for(Swain and Guttmann, 1983):

1. predicting human error probabilities, and
2. evaluating the degradation to the system likely to be caused by human errors alone or in connection with equipment, procedures, or other system and human characteristics that influence behaviour.

The method uses conventional reliability technology with modifications appropriate to the greater variability, unpredictability, and interdependence of human performance as compared with that of equipment performance.

The THERP philosophy is that probabilities of various operations or actions which make up a given task are multiplied together to produce overall system reliability. The task element failures are linked in the form of an *"Human Reliability Event Tree"* (see Figure 6.3), which also includes *"Error Recovery"*. The THERP philosophy also allows *"Human Error Probabilities"* (HEPs) to be modified by various *"Performance Shaping Factors"* (PSFs) such as stress and experience. In addition to PSFs, THERP also accounts for "Dependences" between tasks and between operators.

The method depends heavily on task analysis. The system or subsystem failure that is to be evaluated is defined, after which all human operations involved in the failure and their relationship to system tasks are identified by drawing them in the form of a human reliability event tree. Error probabilities for both correct and incorrect performance of each branch of the Event Tree are predicted by drawing upon a variety of data sources (e.g., Air Data Store, test reports, psychological studies, expert opinion) for inputs. Where an error probability is excessively high, the system is analysed to determine the factors causing that error probability. Changes are then recommended.

Unless specifically stated otherwise all of the HEPs estimates in the Handbook are based on a set of common assumptions that limit or restrict the use of the data as stated. These data apply to situations in which the following hold true:

- The operator's stress level is optimal.
- No protective clothing is worn.
- The level of administrative control is average for the industry.
- The personnel are qualified and experienced.
- The environment in the control room is not adverse.

With regards to mathematics, THERP employs two primary measures (note that the symbology used below is an aid only. Any symbology can be used in THERP):

(1) The probability that an operation will lead to an error of class i(P_i), and

(2) the probability that an error or class of errors will result in system failure (F_i).

P_i is based on what is termed an *human error probability* (HEP), which is the probability that when a given task is performed, an error will occur.

$1 - p_i$ is the probability that an operation will be performed successfully, without error. $F_i P_i$ is the joint probability that an error will occur in an operation and that error will lead to system failure.

$1 - F_i P_i$ is the probability that an operation will be performed that does not lead to error and system failure.

$Q_i = 1 - (1 - F_i P_i)^{n_i}$ is the probability of a failure condition existing as a result of class i errors containing n_i (independent) operations.

Total system or subsystem failure probability resulting from human error is expressed as:

$$Q_t = 1 - [P_i{}^{n}{}_{k=1} (1 - Q_k)]$$

where Q_t is the probability that one or more failure conditions will result from errors in at least one of n class, and the quantity in brackets is $(1 - Q_1)(1 - Q_2)...(1 - Q_n)$.

7.2.3 THERP Procedure

The general approach used for HRA in PRAs has been to identify, analyse, and to estimate HEPs for human tasks that system and human reliability analysts determine could have major impact on system criteria of interest. Figure 7.1 shows the general form of an HRA that has been used by Sandia National Laboratories human reliability analysts as part of the Interim Reliability Evaluation Program (IREP) and described in Bell (1983).

The steps for conducting a HRA using THERP have been stated by Bell and Swain (1981). Figure 7.2 shows a block diagram of one possible ordering of the basic steps of HRA as specified by THERP. This method requires the analyst to:

(1) Visit the plant, survey the control room, interview the operators.

(2) Review information available from systems analysts about critical operator interactions with plant systems.

(3) Talk or walk through various critical procedures step-by-step with a trained operator in the control room or a simulator or a mock-up.

(4) Do a task analysis for various critical situations, formally listing, diagramming and inter-relating task components on paper.

(5) Develop critical Event Trees.

(6) Assign from tabled values appropriate nominal HEPs for component events.

(7) Estimate the relative effects of PSFs and adjust HEPs.

(8) Assess dependence factors and adjust HEPs.

(9) Determine success and failure probabilities for whole sequences of events, neglecting recovery factors.

(10) Determine effects of Recovery Factors (RFs).
Perform a sensitivity analysis, if warranted.
Supply results to system analysts.

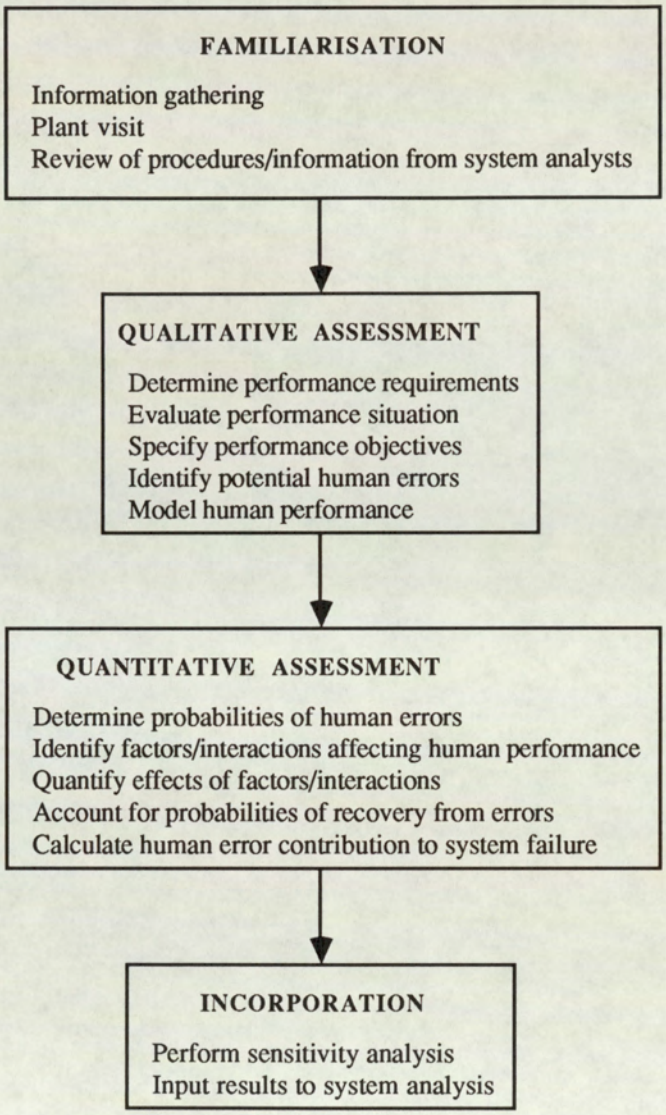


Figure 7.1 The Overall THERP Approach to Performing HRA.
(From Swain and Guttmann, 1983.)

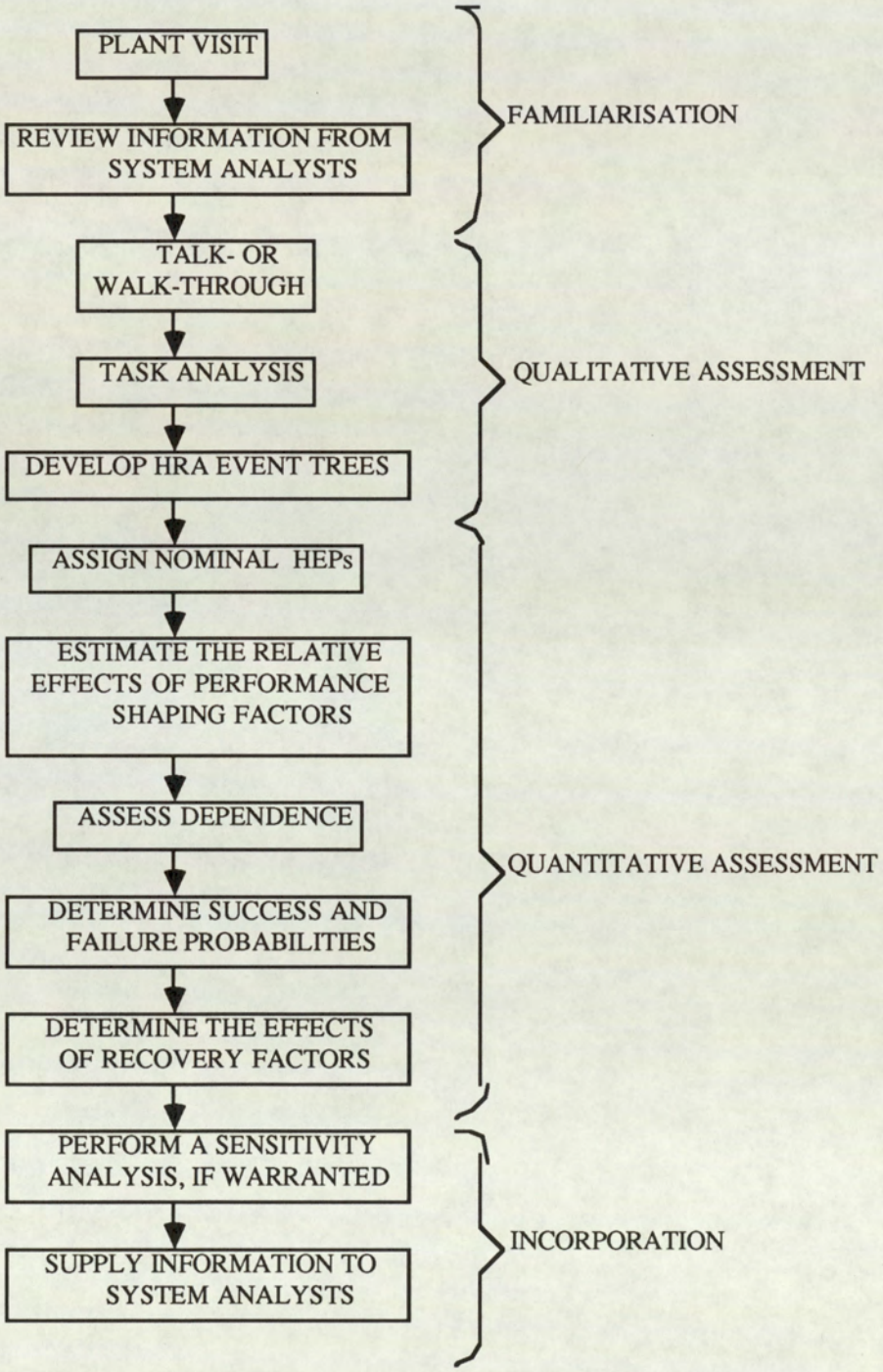


Figure 7.2 Outline of a THERP procedure for HRA
(From Swain and Guttmann, 1983)

Steps 1, 2 and 3 - Plant Visit, Information Review, Walk-Through

The analysis of any situation should be preceded by a visit to the plant. A review of information from system analysts and a walk through the plant are designed to familiarize the analyst with the equipment, procedures, and tasks involved in the analysis. The actual HRA procedure picks up from that point with task analysis.

Step 4 - Task Analysis

Task analysis involves breaking down each task into individual units of behaviour (steps) for which potential errors are identified. This information is entered on a task analysis table, the precise format of which Bell and Swain (1981), consider relatively unimportant. It should include, however, information about the equipment on which an action is performed, the action required of the operator, the limits of his performance, the locations of the controls and displays, and potential errors. The level of detail necessary in the task analysis and the amount of information recorded are determined judgementally. The guiding rule is that one should be able later to recapitulate the rationale for the human error probability (HEP) estimates that were used in the analysis.

Once the individual tasks are identified, the errors likely to be made must be identified for each action/step. A human action (or its absence) constitutes an error only if it has at least the potential for reducing the probability of some desired system event or condition. The determination of the specific errors -omission and commission- must be based on the relevant PSFs and on the task analysis itself. The steps should be listed sequentially.

The determination of potential errors is obviously highly judgemental. Bell and Swain (1981), suggest that extreme care should be exercised in deciding which errors, if any, are to be completely discounted for an analysis.

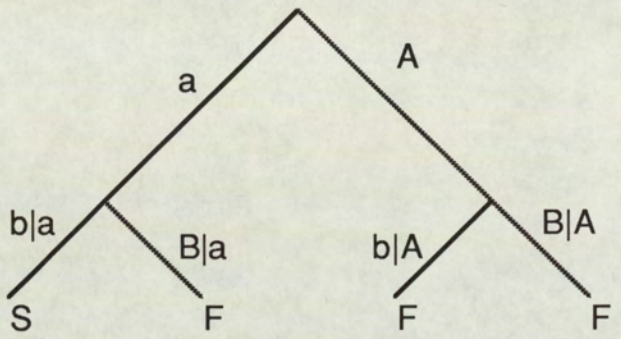
Once potential errors have been identified, the analyst considers other factors that may influence performance but that do not appear in the task analysis.

Step 5 - The HRA Event Tree

The basic tool of THERP is a form of an event tree called the HRA event tree. A schematic example of an HRA-ET is shown in figure 7.3. Each error defined in the task analysis as likely is entered sequentially (in the chronological order in which it might potentially occur if such order is relevant) as the right limb in a binary branch of the HRA-ET. The first potential error starts from the highest point on the tree at the top of the page. Solid lines represent success; dashed lines, error.

Any given task appears as a two-limb branch, with each left limb representing the probability of success and each right limb representing the probability of failure. Thus, at every binary branching, the probabilities of the task must sum to 1.0. Each limb is described by a letter. Capital letters represent the probability of failure of that task. Lower case letters represent the probability of success. The letters S and F represent system success and failure, respectively. In actual practice short descriptions of the tasks or steps are provided along with the symbols.

When the analyst wants to know the probability of all tasks being performed without error, a complete success path through the event tree is followed. Once an error has been made on any task, the system may be presumed to have failed. However, in actuality errors are recovered and probabilities of event success do follow a failure and end in system success.



TASK A = THE FIRST TASK

TASK B = THE SECOND TASK

a = PROBABILITY OF SUCCESSFUL PERFORMANCE OF TASK A

A = PROBABILITY OF SUCCESSFUL PERFORMANCE OF TASK A

b|a = PROBABILITY OF SUCCESSFUL PERFORMANCE OF TASK B GIVEN a

B|a = PROBABILITY OF UNSUCCESSFUL PERFORMANCE OF TASK B GIVEN a

b|A = PROBABILITY OF SUCCESSFUL PERFORMANCE OF TASK B GIVEN A

B|A = PROBABILITY OF UNSUCCESSFUL PERFORMANCE OF TASK B GIVEN A

$Pr[S] = a(b|a)$

$Pr[F] = 1 - a(b|a) = a(B|a) + A(b|A) + A(B|A)$

Figure 7.3 An example of HRA Event Tree Diagramming

Development of the HRA-ET is the most critical (and tedious) part of THERP procedure. If the task analysis has listed the potential errors in order of their anticipated occurrence, the transfer of this information on to the HRA-ET is made much easier.

Step 6 - Nominal Human Error Probabilities (HEPs)

Now that the errors have been identified, defined, and diagrammed, estimates of the probability of occurrence for each must be assigned using the data tables in Chapter 20 of the Handbook or from other data sources. The tables are organized to contain groups of HEPs (and their uncertainty bounds) describing a particular type of error that may occur in the performance of a specified type of task. The description that most closely approximates the error being estimated should be identified. If the differences between the scenario described by the HEP and the scenario being analysed are sufficiently great, the HEP may be used as it is or may be modified to reflect the conditions of actual task performance.

Step 7 - Effects of Performance Shaping Factors (PSFs)

A primary consideration in conducting an HRA is the variability of human performance. This variability occurs within any given individual and also results from the performance of different personnel. Variability is caused by PSFs acting within the individual or on the environment in which the task is performed. Because of this variability, the reliability of human performance usually is not predicted solely as a point estimate but is considered to lie within a range of uncertainty. However, a point value HEP can be estimated by considering the effects of relevant PSFs for the task.

Nominal HEPs are to be used when the scenario outlined in the Handbook reflects the error being analysed. If the analyst judges that the situation under study is more likely to result in error than the one outlined in the Handbook, an HEP closer to the upper

bound than the nominal value should be used. If a plant's situation is judged to be less likely to result in a human error than the one described in the Handbook, an HEP closer to the lower bound than the nominal should be used.

Next, the analyst should consider the influence of PSFs that have a global effect, those that affect the probability of error on all or most of the events in the analysis. The most commonly encountered ones deal with stress and the level of operator experience.

Step 8 - Effects of Dependence

Except for the first branch of an HRA-ET, all branches represent conditional probabilities (probability of a task given success or failure on other tasks) of success and failure. Dependence between events directly affects these probabilities. For any given situation, different levels of dependence (low, medium, high) may exist between pairs of tasks or the performance of two (or more) operators. The Handbook provides guidelines for determining these levels of dependence.

Step 9 - Success/Failure Probabilities

Once errors have been identified and individually quantified, their contribution to the probabilities of system success and failure must be determined, based on the criteria for system success and failure supplied by the system analyst. Multiplying the probabilities assigned to each limb in a success or failure path provides a set of success and failure probabilities that can be then combined to determine the total system success and failure probabilities.

Step 10 - Effects of Recovery Factors (RF)

The term "Recovery Factor" describes the operator's recognition that the error has been made and his repetition of the task once more- this time correctly (Meister, 1985). In performing the HRA the analyst must factor into the analysis the probability that an error will be recovered. Otherwise the probability of system failure resulting from error may be unrealistically high.

Sensitivity Analysis, if Warranted

During the HRA the analyst may wish to determine the effects of manipulating the value of one or more parameters analysed on the total system success probability. The resulting values are then compared to judge the impacts of different magnitudes of changes. This is not a necessary part of an HRA, but is extremely helpful in identifying those elements of the system that have relatively large or small effects on system performance.

Information to System Analyst

The human reliability analyst should present a copy of each HRA-ET along with a synopsis of the results, a copy of the task analysis table, and a list of the assumptions, to the fault tree analyst. Both analysts should then go over the HRA-ET and its associated assumptions very carefully to ensure that the system success has been correctly defined and that the results of the HRA-ET are not applied outside the scope of its stated limitations.

7.2.4 THERP Advantages

(i) THERP can be used easily for design, risk and reliability assessments at all stages of system development. The level of detail of the THERP analysis can be tailored according to the depth of the overall assessment being performed.

(ii) THERP can be integrated into PRA easily. Its form and approach make it compatible with fault tree methodologies.

(iii) THERP provides a structural, logical, well documented record of the factors and errors considered in the HRA. This allows the results to be reviewed easily and assumptions used to be examined.

(iv) THERP is the most used, particularly in the nuclear industry, and acceptable technique to regulatory bodies and scientific community. It is the method currently being recommended by the US Nuclear Regulatory Commission (NUREG/CR-2300, 1981). and is high on 'face validity'.

7.2.5 THERP Disadvantages

A number of factors complicate the THERP procedure:

(i) A good deal of judgemental fine-tuning is involved in the selection of THERP estimates. In order to implement the procedure it is necessary to determine all possible errors, select error probabilities appropriate to the anticipated errors that have been identified, determine the degree of dependence/independence among tasks and operators, and determine the PSFs that affect error likelihood.

(ii) Each of the above tasks requires much judgement, an in-depth (human factors/ mathematical, etc.) knowledge and high experience. For example, the criteria for selecting the appropriate error probabilities from the data sources are not precisely specified. The data sources THERP uses generally provide a range of error values with a normative or nominal value. Depending on the PSF operative in the error situation and the degree of dependence among tasks, the correct error value shifts up or down in the range of values provided. To select the correct error value from the data source requires tedious calculations, expert judgement and considerable practice.

(iii) In addition to the above disadvantages, the data bank and human performance models provided by THERP are not well suited to analyse and predict all human (particularly cognitive) activities and in various (particularly critical) situations of interest. However, an attempt has been made in the Handbook to extend the behaviouristic approach THERP to cover cognitive activities.

(iv) THERP requires data concerning the task (e.g. detailed procedures etc.), and if these do not exist the analyst must make assumptions. Although the user can learn how to apply the technique from the THERP handbook, this would take a long time.

7.3 SLIM Method

7.3.1 SLIM Background

Embrey (1981) has suggested a methodology similar in some respects to that used in developing the Likelihood of Accomplishment Scale used in TEPPS -The Establishment of Personnel Performance Standard (Blanchard et al., 1966).

The theoretical basis of the Embrey approach is what is known in decision analysis as Simple-Multi-Attribute Rating Technique or SMART (Edwards, 1977), which in Embrey's approach is called the Success Likelihood Index Method (SLIM). SLIM has been evolved by Embrey as part of a general approach to the evaluation of human reliability in systems. The impetus for SLIM (as for all the expert judgement techniques) is the lack of objective error data.

Several evaluations of SLIM have been performed by its originator. A pilot evaluation of SLIM has performed by Embrey (1983) to determine its practical feasibility. Initial results appear to be encouraging, although, as Embrey concludes, there are difficulties (e.g., inconsistency among judges) that must be overcome. The SLIM technique is still in the development stage.

7.3.2 SLIM Description

SLIM is a systematic method for positioning the likelihood of success of a task on a scale as function of the various conditions affecting successful task completion. The rational is that the likelihood of an error occurring in a particular situation depends on the combined effects of a relatively small set of PSFs. PSFs are the various aspects of the task, the individual or the environment which affect performance. The absolute probability of success for tasks placed on this scale can be determined by calibrating the scale with reference tasks for which success probability is known.

7.3.3 SLIM Procedure

The operation of SLIM is a consensus process (Embrey, 1983). The recommended procedure for using SLIM involves the following steps (Embrey et al., 1984):

Step 1 - Modelling and Specification of PSFs

This step involves the definition of the situation to be studied. It includes collection of information regarding the characteristics of the tasks, the individuals who will perform them, and determination of the most relevant PSFs and their relative importance to the tasks under consideration. The tasks are then subdivided into subsets for which success probabilities are influenced by common sets of PSFs.

Step 2 - Weighting of PSFs (see Table 7.1 for an example)

The PSFs are weighted by judging the importance of each factor in terms of its likely effects in either enhancing or degrading the reliability (probability of success) of the specific task.

Table 7.1 PSFs Weights
(from Embrey et al., 1984)

PSF	Assigned Weight	Normalized Weight	
Quality of Information	100	100/200	0.50
Training	50	50/200	0.25
Time Available	30	30/200	0.15
Procedures	20	20/200	0.10
	Total = 200		Total = 1.00

Step 3 - Rating the task

A numerical value (rate) is assigned to each PSF (usually between 1 - 100) to reflect its actual quality (influence) in enhancing or degrading reliability of the specific task. The ratings on a particular PSF are relative to the ratings for all the other tasks on that PSF.

Step 4 - Calculation of SLI (Success Likelihood Index).

The SLIs for each task are calculated using the formula shown below:

$$SLI_j = \Sigma^{x}(R_{ij}.W_i)$$

where:

SLI_j = the Success Likelihood Index for task j (j = no. of tasks)

W_i = the normalized importance weight for the i^{th} PSF (the weights for all the PSF sum to 1).

R_{ij} = the scaled rating (position on the scale) of the j^{th} task on the i^{th} PSF.

Table 6.2 shows an example of SLI calculation in a SLIM session.

(a) The Product of the Weightings and Ratings for each PSF represents the relative effect of each dimension.

(b) The sum of the products is the index of the overall effect of the PSFs on Human Reliability.

Step 5 - Conversion of the SLI to Probability

This step involves the substitution of (at least) two known task probabilities into the equation:

Log (probability of success) = a SLI + b.

where *a* and *b* are constants.

Table 7.2 Calculation of the SLI (from Embrey et al., 1984)

PSF	Normalized Weights (From Table 8.2)	Rating	Product Weight x Rating
Qual.of Info.	0.50	70	35.0
Training	0.25	20	5.0
Time Avail.	0.15	10	1.0
Procedures	0.15	50	5.0
			SLI = 46.5

Determination of the constants in the equation requires that at least two tasks for which the HEPs are known are included in the SLIM session, and that the SLIs for these tasks are assessed.

7.3.4 MAUD Description

SLIM has been implemented through the use of an interactive computer program called MAUD (Multi-Attribute Utility Decomposition). The MAUD software was developed by Dr P. Humphreys and Ms. A. Wisudha of the London School of Economics. MAUD is used for eliciting from judges the ratings and weightings of PSFs. An example and a detailed description of the technique are given in Embrey et al. (1984).

With MAUD, judges first rate PSFs and then weight them, thus reversing the order used in SLIM procedure.

A separate computer program is used to convert the SLI values into probabilities using the calibration equation derived from the two reference tasks (see Step 5).

7.3.5 SLIM Advantages

(i) A major advantage of SLIM approach is that it explicitly identifies the PSFs which are judged to be major determinants of the probability of error in the tasks being assessed. The weights which are assigned to these factors can be used to provide design recommendations with regard to which changes will have the greatest effect in reducing the likelihood of error.

(ii) Another advantage of the approach is that it is highly scrutable, i.e., the means via which the final result is arrived at is accessable to external auditing and review. In other subjective techniques, such as APJ, or Paired Comparisons (PC), this is not possible since the process via which the judges arrive at their conclusions is covert.

(iii) SLIM does not require extensive decomposition of the task considered to smaller units or subtasks (as does THERP), but it tends to take a more holistic/global approach to the description of tasks and the quantification of human reliability.

(iv) SLIM is more suited for handling cognitive task elements (than for example, THERP).

7.3.6 SLIM Disadvantages

(i) The first major disadvantage of the SLIM method is that it makes extensive use of expert judgement requiring a multidisciplinary team of experts (eight experts, for example, were used in the exercise carried out to test SLIM-MAUD, described in Rosa et al., 1985).

(ii) The personnel and resources required for setting up SLIM-MAUD data bases are greater than with some techniques.

(iii) Another major disadvantage of SLIM is that it requires the availability of two known probabilities of failure and requires the use of a logarithmic calibration equation to derive probabilities for the set of tasks considered.

(iv) The technique also requires homogeneity of the events or tasks to be analysed. If the tasks are not homogeneous the validity of the results may be affected.

7.4 APJ Method

7.4.1 APJ Background

Absolute Probability Judgement (APJ), also known as Direct Numerical Estimation, (Comer et al., 1984), is the most direct approach to the quantification of HEPs. It relies on the utilization of experts to estimate HEPs, based on their knowledge and experience.

The main reason for using APJ for estimating human error probabilities is (as for SLIM and all the other judgement techniques) the lack of relevant or useful objective human probability data.

The rationale for using APJ is that there exist experts who have experience and/or knowledge which can be translated into quantitative estimates of the probability of occurrence of an event.

The two main requirements for using APJ are that :

i) The expert must be familiar with the type of the problem assessed.

ii) The expert must be able to accurately translate this expertise into probabilities,i.e., possess sufficient knowledge of the calculus probabilities.

There are two forms of APJ, namely Group APJ methods, and the Single Expert Method, (referred to as engineering judgement).

Much research involves the use of group methods, since in many circumstances it is unlikely that a single expert has enough relevant information and expertise to accurately estimate human reliability. In the group methods, therefore, individuals' knowledge and opinions are aggregated either mathematically or by bringing the judges together to arrive at some form of consensus agreement. Engineering judgement, on the other hand, refers to a single expert estimating a HEP. This latter approach is frequently used both in HRA and in hardware reliability assessment, though it is arguably a less reliable approach.

7.4.2 Group APJ Methods Description

The APJ methods described here are group methods, since these are preferred, as they are less prone to the biases and knowledge limitations of a single expert.

There are four major group methods, each briefly described below (see UKAEA, 1988).

1. Aggregated Individual Method

This method does not require that the experts meet, but they make estimates individually. These estimates are then aggregated statistically by taking the geometric mean of the individuals' estimates for each task.

The main disadvantage of this method is that the experts do not share their expertise.Its main advantage is that it does not require experts to be co-located, therefore, it avoids personality conflicts which may bias experts' probability estimates

2. Delphi Method

In the Delphi method (Dalkey, 1969) the experts (as in the individual method) do not meet. The experts make individual reliability assessment, but in this case all HEPs are shown to all experts. The individual experts then reassess the HEPs they have previously estimated. These are then statistically aggregated.

Although this method allows information to be shared, it still does not allow experts to discuss and resolve their different perspectives.

3. Nominal Group Technique (NGT)

The NGT is similar to the Delphi method except that some limited discussion is allowed between experts for clarification purposes. The assessments are then statistically aggregated.

This method may be considered as an improvement over the first two methods since it allows and enhances sharing of information.

4. Consensus Group Method.

In this method each expert contributes to the discussion to arrive at an estimate upon which all members of the group agree.

This method maximizes information sharing but necessitates expert co-location. In that case personality conflicts may arise affecting the experts estimations.

7.4.3 APJ Procedure

This subsection shows how to carry out a group APJ exercise. The overall procedure is as follows (see Comer et al., 1984, UKAEA, 1988):

1. Selection of subject matter experts
2. Preparation of task statements
3. Preparation of response booklets
4. Development of instructions for subjects
5. Obtention of judgements
6. Calculation of inter-judge consistency
7. Aggregation of individual estimates
8. Estimation of uncertainty bounds.

Step 1 - Selection of Subject Matter Experts

The experts selected for making judgements must be familiar with the plant/system and tasks to be assessed. Seaver and Stillwell (1983) suggest that six experts would be sufficient, although more would be preferable. In practice however, a smaller group of three to four experts is often used.

Step 2 - Preparation of Task Statements

This is a critical aspect of the APJ procedure. The more fully the tasks are described, and the assumptions (PSFs) defined, the less they will be open to individual interpretation by the experts. Instructions in form of diagrams, etc., may also be useful.

Step 3 - Preparation of Response Booklets

Preparation for response booklets is preceded by selection or design of the scale on which the experts will indicate their judgements (examples are described in Seaver et al., 1982).

The instructions, assumptions for the task set, and sample items included in the response booklet should appear first (examples are given in Seaver et al., op cit.). Then, if a consensus group is not used, the tasks are presented in random order to minimize the effects of task presentation sequence.

Step 4 - Development of Instructions for Experts

The instructions given to experts at the beginning of a session should indicate the purpose of, and reasons for the study. Also instructions on how to elicit uncertainty bounds must be given to experts.

Step 5 - Obtention of Judgements

Experts are asked either to work through their booklets, or in a group consensus mode, to discuss each task in turn and arrive at a consensus estimate.

If a consensus group is being utilized, a facilitator will be required to overcome any personality/group problems and to prevent any biases from negatively affecting the judgements.

Step 6 - Calculation of Inter-judges Consistency

The following procedure for calculating inter-judge consistency (UKAEA, 1988) is based on Seaver and Stillwell, 1983, and uses the Analysis of Variance (ANOVA) technique (Meddis, 1973; Chatfield, 1978). An example set of HEPs are shown in Table 7.3

1. Calculate the column totals (t): e.g. -6.75, -7.53, etc.
2. Calculate the row totals (r): e.g. -6.82, -2.83 etc.
3. Calculate the grand total (T): T = -32.35
4. Calculate the correction term (C): $C = T^2/mxn$

   $C = (32.35)^2 /16 = 65.41$
5. Calculate the sum of squares of raw scores: $\sum x^2$

   e.g.$(-1.9)^2 = ...(-2.60)^2 = 83.05$
6. Calculate the total sum of squares (TSS) :$\sum x^2 - C$

   = 83.05 - 65.41 = 17.64
7. Calculate the 'Between Column Sum of Squares' (Col SS): $= \frac{\sum t^2}{n} - C$

$$\text{Column SS} = \frac{(-6.75)^2 + (-7.53)^2 + (-9.62)^2 + (-8.45)^2}{4} - 65.41$$

= 66.55 - 65.41
= 1.14

Table 7.3 APJ-derived HEPs (Kirwan, 1982)

Expert (m)	1	2	3	4
1	0.013	0.07	0.0025	0.068
2	0.56	0.13	0.33	0.063
3	0.005	0.01	0.00017	0.00033
4	0.005	0.00033	0.0017	0.0025

The set of HEPs obtained from the experts are first translated into their logarithmic equivalents. The results are shown in Table 7.4 below.

Table 7.4 Log HEPs

Expert (m)	1	2	3	4	Total
1	-1.90	-1.15	-2.60	-6.82	-6.82
2	-0.25	-0.90	-0.48	-2.83	-2.83
3	-2.30	-2.00	-3.77	-11.55	-11.55
4	-2.30	-3.48	-2.77	-11.15	-11.15
Total	-6.75	-7.53	-9.62	-8.45	-32.35
Average	-1.69	-1.88	-2.41	-2.11	

8. Calculate the 'Between Row Sum of Squares'

$$\text{(Row ss)}: = \frac{\sum r^2}{m} - C$$

$$\text{Row SS} = \frac{(-6.82)^2 + (-2.83)^2 + (-11.55)^2 + (-11.15)^2}{4} - 65.41$$

$$= 78.06 - 65.41 = 12.65$$

9. Calculate the 'Residual Sum of Squares':

   Residual SS = TSS - Col SS - Row SS

   = 17.64 - 1.14 - 12.65 = 3.85

10. Enter the appropriate degrees of freedom into the summary table (Table 7.5)

Columns differential df	= Number of columns -1	= 3
Rows differential df	= Number of rows -1	= 3
Total differential df	= Number of scores -1	= 15
Residual differential df	= Total df - Col df - Row df	= 9

Table 7.5 Summary ANOVA table

Source	Sums of Squares	df	Variance	F-ratio
Events (Columns)	1.14	3	0.38	0.88
Judges (Rows)	12.65	3	4.22	4.22
Residual (Error)	3.85	9	0.43	
Total	17.64	15		

11. Calculate the variance estimates by dividing each of the sums of squares by the appropriate degrees of freedom. Therefore:

Column (events) Variance	= 1.14/3	= 0.38
Row (expert) Variance	= 12.65/3	= 4.22
Residual Variance	= 3.85/9	= 0.43

12. Calculate the F-ratios

    F (columns) = (Column Variance) / (Residual Variance)

    F (rows) = (Row Variance) / (Residual Variance)

Therefore:

F (columns) = 0.38/0.43 =0.88

F (rows) = 4.22/0.43 =9.81

13. The last step is to determine the interclass correlation co-efficient according to the following formula:

$$r = \frac{F - 1}{F + (n-1)} = \frac{0.88 - 1}{0.88 + 3} = -0.03$$

Step 7 - Aggregation of Individual Estimates

If a consensus group is not used, and agreement between judges is adequate, it will next be necessary to aggregate individuals estimates for each HEP by taking the geometric mean of the individual estimates.

Step 8 - Uncertainty Bounds (UBs)Estimation

UBs are calculated using the following formulae (Seaver and Stillwell, 1983).

Log HEP $\pm$ 2 s.e.

where s.e. = standard error $= \sqrt{\underset{m}{V}(\log \underline{HEP_i})}$

If using a group consensus, UBs can be estimated using the APJ m,method itself. Each expert is asked to estimate UBs as detailed in Step 4. These can then be aggregated statistically.

7.4.4 APJ Advantages

The APJ technique has been shown to provide accurate estimates in fields other than human reliability assessment (e.g., "weather forecasting", see Murphy and Winkler, 1974). Other studies (such as "Service Sector", Williams, 1983; "Off-shore Drilling", Bellamy, 1985, and "Experimental Study", Kirwan, 1982) also give support for the validity of this method.

7.4.5 APJ Disadvantages

The main disadvantage of the APJ method are:

(i) The technique relies heavily on expert judgements and, therefore, is prone to many biases, such as overconfidence or conservatism (for further details, see Kahneman and Tversky, 1979), and to personality/group problems and conflicts, which can, if not effectively countered (e.g., by the use of a facilitator) significantly affect the validity of the technique.

(ii) APJ has relatively high resource requirements, due to its use of multiple experts. Experts using APJ are also required to have substantive expertise,i.e. familiarity with the problem, and normative expertise, i.e. expertise in statistics to enable them to translate their expertise into probabilities. Knowledgeable experts may be difficult to obtain, therefore training may be required.

(ii) APJ has virtually no sensitivity analysis capability.

7.5 CONCLUSION

The results of the evaluation of THERP, SLIM, and APJ with respect to the specific criteria described in Chapter 6, have been discussed in this chapter and can be summarized as follows:

(a) Both THERP and SLIM techniques make use of task analysis to break down higher-order (more molar) operations into tasks.

(b) All techniques examined require in-depth knowledge and high experience and, hence, cannot be usable by non-experts.

(c) All techniques require tedious calculations.

(d) All techniques make use of expert judgements to estimate human reliability data not readily available.

(e) All techniques can be used at all stages of system development. They can also be employed at any level of behaviour (molar or molecular).

(f) Both THERP and SLIM techniques provide design recommendations and sensitivity analysis capabilities.

(g) Although THERP makes less use of judgements than SLIM and APJ, such judgements are much informal than in SLIM.

(h) SLIM and APJ tend to take a more holistic approach to the evaluation of a task than THERP.

(i) SLIM and APJ are more suited for handling cognitive task elements. However, an attempt has also been made in the "Handbook" to extend the THERP approach to cover cognitive activities.

(j) Among the three techniques, APJ is the technique which relies more on expert judgements and expertise, which may be biased, or not comprehensive enough to generate accurate human reliability data. APJ is also the least sophisticated and reliable of the three selected HRA techniques.

(k) SLIM is the most resources (person-hours) demanding.

(l) None of the methods satisfies all the criteria described in Chapter 6, but both THERP and SLIM come within "nodding distance" of them.

It is clear from the discussions presented above and in Chapters 5 and 6, that HRA is a complex process, which requires a familiarity with many areas, such as psychology and ergonomics, which are not generally considered to fall within the engineering and management disciplines. A large level of judgement and effort is required to select and perform an appropriate HRA which can only be acquired by extensive and costly training. In addition to that, the number of human factors experts in the HRA domain is handful.

Clearly, these factors (among many others not mentioned here) have greatly affected and limited the widespread use of conventional HRA methods by non-experts.

It is apparent, however, that the effective assessment of the reliability of a system cannot be carried out without an adequate consideration of the human element.

Therefore, tools are needed to assist the non-expert, as well as the expert, in the selection and detailed application of the appropriate HRA to the specific situation considered. The expert system developed here, which is based on THERP, SLIM, and APJ techniques, is an attempt to overcome some of the HRA deficiencies and is described in the next chapter.

Although no one of the three selected HRA techniques satisfies all the desirable requirements discussed in the previous chapter, it was thought that the combination of the capabilities of more than one technique would benefit the user in assessing human reliability.

CHAPTER 8

DEVELOPMENT of HERAX

Expert System Approach for Human Reliability Analysis

8.1 Introduction

The survey of expert system technology provided in Chapter 2 shows that expert systems have been applied to problems complex and involving the use of incomplete and improper data with many kinds of uncertainties, such that a substantial amount of human judgement is needed in their solutions.

It is also clear from the discussions in the previous chapters, that existing human reliability assessment methods are numerous, complex and require a high level of subjective judgement and experience for their selection and application. These characteristics make the HRA problem domain ideal for expert system technology.

Therefore, the use of Artificial Intelligence (AI) techniques for the development of an expert system to aid non-expert in the selection and systematic application of HRA procedures, as part of probabilistic risk assessment (PRA) studies in nuclear and process plants, is worthwhile.

This chapter provides a detailed description of the structure and methodology of an expert system that has been developed as a first attempt to achieve that purpose.

8.2 HRA Experts' Problem-Solving Tasks

The nature of the human reliability analysis experts' knowledge and reasoning processes involved in carrying out the different HRA procedure's steps have been discussed in detail in Chapters 5, 6 and 7. These types of knowledge and the corresponding methods of reasoning are briefly considered here.

The HRA procedure's steps and the corresponding cognitive, problem-solving tasks that are faced by a HRA expert when analysing any particular case involving human actions are many and varie from one HRA method to another. However these can be viewed as consisting of two main stages or two classes of knowledge leading to different types of reasoning models:

(i) In the initial stage the task of the HRA expert is one of *qualitative analysis* or *modelling*, i.e., the analyst *collect* information, *define* the situation, *identifies* the likely human errors and *classifies* them with reference to a particular task or human error taxonomy, such as "omission" and "commission". The modelling or qualitative analysis task involves also, the *identification* of the Major Performance Shaping factors (PSFs) and *decisions* about their importance and effects on enhancing or degrading human performance. Finally, consequences of human errors on the system performance is *estimated.*

(ii) Once a modelling of human errors is made, *quantification* or *predictive problem solving* is the next HRA expert's task. This task involves human error probabilities (HEPs) *assignment*, effects of PSFs, dependences and recovery factors *assessments*, and *deduction* of the possible human errors consequences on system safety and reliability. The final task is a *"what if"* or *sensitivity analysis,* if warranted.

8.3 Why an Expert System for HRA?

Prior to the description of the architecture and features of the proposed expert system, it will be instructive to first discuss the need, rational or motivations, for using AI and expert system approaches for this problem.

8.3.1 HRA's Knowledge Deficiencies

1. The discussions of the nature of the problem domain carried out in Chapters 5 and 6, show clearly that, whereas human reliability is critical to the overall safety and reliability of process plants, HRA problems are complex, and thus require knowledge-intensive problem solving, i.e., in solving these problems, HRA practitioners rely heavily on knowledge and experience accumulated over a period of many years.

2. In addition, the domain of Human Reliability Analysis is vast, continually changing, poorly formalized, and not readily available, thus needs to be used selectively.

3. Even with the availability of the relevant knowledge, the supply of qualified ergonomists/human factors experts in HRA is limited and expensive. Consequently, the contribution of human performance to plants safety and reliability will be inevitably analysed by non-experts.

4. The conduction of the existing HRA procedures involves a series of complicated analytical steps not always clear, and, in addition to being time consuming and tedious particularly for non-experts like engineers and managers, it requires specialized training and thorough knowledge in many disciplines, such as psychology, ergonomics, engineering, and statistics. The acquisition of this knowledge is expensive and takes many months to build.

The above problems are some of the main factors that have limited the widespread use of the existing HRA methods, and hence, limited the systematic analysis of human reliability by engineers and managers when performing overall system reliability and safety studies.

It is obvious, therefore, in view of these shortcomings, and because of the many general advantages of expert systems discussed in Chapter 2, that it is worthwhile to construct an expert system that would overcome some of the difficulties associated with selecting and applying conventional HRA approaches by non-experts.

8.3.2 Advantages of Expert Systems to HRA

Some of the specific features and capabilities that make an expert system potentially useful for HRA and that have motivated this project are described below.

1. The development of an expert system could serve to capture and represent the knowledge and expertise of HRA specialists in a form that can be easily understood and applied, and make that expertise and advice available to those who need it.

2. It could help systematize human reliability analysis procedure, and improve its consistency and standardization. This would encourage the widespread use of HRA techniques by engineers, designers, and managers, or other intended users, when conducting overall system safety and reliability/risk analysis studies.

3. The system could be a good means for pooling the expertise of a number of specialists, to produce a technique that is more effective than any of them working alone.

4. The expert system could lead to a much faster problem solving process, by incorporating many details of the HRA procedures into the inference engine module and removing them from the user's responsibilities.

5. In addition to alleviating the problems of lack of experts in this area, and the complexity of the existing HRA techniques, the system should be able to justify its decisions to the users. With this explanation capability the system could also serve as a training tool.

6. A final reason for advocating the expert system approach in human reliability analysis has to do with the changing nature of human behaviour sciences theories and methods, which constitute the basic foundations of human reliability analysis methods, such as psychology and ergonomics/human factors. In addition to that, process plants are becoming more automated, and more complex, and, consequently, their design, operation and maintenance procedures are continually changing. As a result of these changes, new HRA techniques are constantly being developed, reviewed, and updated. A computer-based system could play an important role in helping to bring this evolving knowledge to the non-expert users in a more organized and simplified way. It was also felt that a system such as the one proposed here could be altered and its knowledge base extended much more easily if it were structured as a rule-based rather than in a more conventional format.

8.4 Knowledge Source and Acquisition Procedure

There are two major sources of knowledge for expert systems developments. These are the literature, and human experts.

In discussing the advantages and disadvantages of both techniques, Taylor et al., (1987), state that:

> *"In general, the literature on a subject will be more structured than a transcript compiled from an interview with a human expert. it will also allow the knowledge engineer to acquaint himself with the basics of the domain without wasting the time of an expert. The literature is therefore a more suitable starting point for the knowledge acquisition process. The main drawbacks of the literature are that it is not interactive and it may well be incomplete...*(in that case) *the knowledge engineer will have to approach human experts."*

However,

> *"If the literature is sufficiently detailed and complete (...) then a human expert may not be needed until the validation stage."*

The approach applied to acquire the knowledge and data necessary for the building of the proposed expert system for HRA was based mainly on various careful reviews and analyses of the available and published literature.

The following methodology was used:

First, published human error and human reliability literatures as well as related disciplines, such as human performance modelling and evaluation, man-machine system/interface, and hazard and risk assessments, were reviewed to gather the information needed to develop the expert system. The survey conducted covered more than 30 years of human reliability development. The knowledge gathered during this literature review comes from various sources. Included among these are the following:-

1. Case studies and solved example problems.

2. Handbooks, reference books and guides.
3. Unclassified reports (in paper and microfiche formats).
4. National and international conference proceedings.
5. Review articles and several hundreds of scientific papers and articles in journals and magazines.

Second, these knowledge sources were classified and analysed to determine how relevant and useful they are for providing the knowledge needed for the HRA expert system development and, therefore, worth a further, more detailed analysis.

Third, to the selected, worthwhile literature, "IF-THEN" induction rules were applied to extract the information necessary to develop the knowledge base, the inference engine and the explanation modules of the expert system.

8.5 Hardware and Software Environments

It was decided from the beginning to develop the expert system for HRA from scratch, i.e., using a computer programming language environment, instead of using a "Shell" environment. Therefore, the first step in implementing the system was to decide on the type of the computer (*hardware*) and the computer programming language (*software*) upon which the system will be constructed. When this project started in October 1985, in the Health and Safety Unit in the (then) Department of Chemical Engineering, an IBM-PC/AT and an implementation of the programming language LISP (LISt Processing or Programming), called GCLISP (Golden Common LISP), have just been acquired, so it was decided to develop the initial prototype of the HRA expert system on this microcomputer using GCLISP.

(i) LISP Features

LISP, the second oldest (developed at MIT in 1959) language still in everyday use, is the major language for work in artificial intelligence. All the large early expert systems were developed in LISP or a tool written in LISP. LISP deals with symbols (as well as numbers). A symbol can be any combination of characters; however, LISP deals primarily with alphanumeric strings that look like words (atoms) or sentences (lists). Atoms and lists collectively are called symbolic expressions.

In an article written in 1978, McCarthy, inventor of LISP, has cited six features that account for LISP's uniqueness among programming languages:

1. Its ability to compute with symbolic expressions in addition to numbers; that is, bit patterns in a computer's memory and registers can be defined to stand for arbitrary symbols, not just those of arithmetic.
2. Its capacity for list processing; that is, representing data as linked-list structures in the machine and as multilevel lists on paper.
3. Its extensibility; Control structure based on the composition of simple functions to form more complex functions.
4. Its powerful use of recursion on all levels, as a way to describe process and problems.
5. Representation of LISP programs internally as linked lists and externally as multilevel (structured) lists, that is, in the same form as all data are represented.
6. The EVAL function, written in LISP itself, serves as an interpreter for LISP and as a formal definition of the language.

In essence, LISP makes no distinction between data and programs, so LISP programs can use other LISP programs as data. This feature has helped a great deal to speed up the expert system programming process. LISP is a highly interactive, flexible, and

recursive language. The major advantage of LISP is its unique "nesting" nature. This lend powerful capabilities to many problem-solving techniques like searching. With recursion, LISP can break problems into smaller problems where the program calls itself with simplified arguments. However, LISP's recursion capabilities must have a termination point -it cannot recur for ever. These properties of LISP do not always make for easily read syntax, but they allow for elegant solutions to complex problems that are very difficult to solve in the various conventional programming languages.

A major reason LISP is so popular is that a LISP program can be naturally represented in LISP data structures. Since programs and data have the same form, they can be interchanged at will, allowing developers to write programs that can, themselves, run and modify other LISP programs. This process lets an expert system program modify lines of its own code while the program runs. Another attribute of LISP is that memory management is completely automatic, and data typing and storage allocation take place at program runtime. LISP relies on dynamic allocation of space for data storage, so the developer does not have to worry about assigning program space. This makes LISP very modular as property lists need not be adjacent in memory since everything in LISP is done with pointers to select and identify needed data. As a result, LISP manages storage space very efficiently, freeing the developer to create more complex flexible programs.

LISP has only a few basic functions. All other LISP functions are defined in terms of these. Thus, the programmer can easily create a LISP operating system and then work up to whatever higher level is desired. Because of this flexibility, until recently LISP was not standardized in the way FORTRAN (the oldest language) and BASIC were, and people often had trouble moving between one dialect of LISP and another. By 1984, however, a standardized version of LISP, COMMON LISP (Steele 1984; Winston and Horn 1984), had replaced all the other dialects in the commercial marketplace.

Other specialized features of LISP include its powerful debugging facilities, the availability of both a compiler and interpreter for program development, its automatic runtime checking, and its macro facility that allows for easy extensions of the language.

LISP can also be used for the development of software. The best example of this is the operating system and software environment for the LISP machines. On most machines, the operating system, interpreters, compilers, editors, and utilities are all written in LISP.

The principal syntactical device used for representing lists in LISP are nested parentheses. Many inexperienced people cite this as one of the weaknesses of the language, complaining that it makes it extremely hard to read and debug. But for an experienced LISP programmer, the parentheses are a very useful aid in displaying a program's structure, and there are utilities to help the programmer avoid syntax errors.

(ii) GCLISP 286 Developer

GCLISP is a subset of COMMON LISP, implemented by Gold Hill Computers (1986) in the USA, and is being marketed in the UK by AI Ltd at Watford. The first version of GCLISP requires PC/MS-DOS Version 3.0 or 3.1, and a minimum of 512K bytes of base memory and 1 megabyte of extended memory.

The source code of the first version of the prototype expert system HERAX was developed using version 1.0 of GCLISP interpreter. The current version of HERAX program has been developed using Gold Hill computers ' GCLISP 286 Developer, which is a LISP development system for the IBM PC AT and 100 percent compatibles. It contains version 2.2 of the interpreter and version 1.0 of the compiler, requires PC-

DOS or MS-DOS version 3.0 or higher, and can use up to 15 megabytes of extended memory.

Like GCLISP version 1 for the IBM PC and compatibles, The GCLISP 286 Developer comes with the GMACS text editor, a debugging utility, an interactive tutorial, on-line help, a GCLISP reference manual, and two books: The Common LISP Reference Manual by Steel (1984) and LISP (second edition) by Winston and Horn (1984).

The interpreter version 2.2 has been enhanced to take advantage of the large address space of the Intel 80286 microprocessor and is called the Large Memory (LM) interpreter. It needs a minimum of 512K bytes base memory (DOS-accessible), and 2 megabytes extended memory in standard IBM PC AT protected mode. The compiler, called the LM Compiler, requires at least 3 megabytes of extended memory and 700 K bytes of space on the hard disk.

GCLISP 286 Developer also supports the Golden Common LISP **RUNTIME**. GCLRUNTIME is a software that creates executable, standalone programs from compiled GCLISP Developer programs. These programs can be invoked directly by the end user from the top level of the DOS operating system environment. So, the end user needs no LISP knowledge to use the application runtime.

8.6 Development Methodology

The development philosophy was to get a prototype system up and running as quickly as possible for early evaluation. The capability of LISP for rapid prototyping has helped to achieve this objective easily. A first attempt was made to develop a simplified, relatively intelligent program for HRA, based on one of the existing techniques, THERP. Although this first attempt contained a number of deficiencies, it served as a catalyst for uncovering a vast wealth of HRA knowledge previously not

thought relevant. Everything in the system was reviewed at that time, resulting in a great number of misconceptions being uncovered and the discovery that there was a large amount of new knowledge still to be acquired.

The system was iteratively enlarged and refined. Each review resulted in refinement of the rules and the user interface.

About halfway through the project it was decided to expand the range of HRA techniques to include a second procedure, based on the subjective method SLIM. A third procedure, based on the direct judgement technique APJ, was added later on.

It has been found from the discussion of the nature of the HRA problem domain in the previous chapters, that the application and selection of conventional HRA techniques require a high degree of judgement and expertise in a number of scientific disciplines, and that there are different types of problem solving involved in the HRA domain of expertise, for example, the problem solving involved in human reliability modelling is of a type different from that involved in reliability quantification. Also, the judgement and expertise involved in each HRA approach are different.

In order to represent the concepts of this complex domain, the analysis of human reliability in complex systems such as process control plants, it was necessary to formulate an architecture where for each type of HRA procedure and for each type of problem solving, there exists in HERAX a separate knowledge structure, with the associated problem-solving mechanisms embedded in it. Thus that specific structure can be viewed as an active knowledge structure for problem solving of that type. Contrast this with the traditional view in which knowledge has an existence independent of the problem solvers that may use it. It was also necessary to develop a framework and apply a set of techniques that allow integration of these different types of reasonings from multiple experts using a single system.

To test the methodology, a hypothetical case study involving runaway reaction in a process plant was used and is presented in Chapter 9.

8.7 HERAX Architecture

Having outlined the main problem-solving tasks performed by HRA experts, discussed the rational for adopting an expert systems approach for human reliability analysis, described the knowledge acquisition procedure, the hardware and software, and finally the methodology used to develop the system, this section will focus now on the description of the structure of the system which was actually developed.

HERAX (*the **H**uman **E**rror and **R**eliability **A**nalysis e**X**pert*) is a rule-based expert system (earlier version of the system was called **RELAS** for *"**R**eliability **A**nalysis **S**ystem"*, and is briefly described in the book published by Kibblewhite, 1988) which incorporates three existing HRA methods. It is designed to be used for the modelling and quantification of human errors/reliability under normal and abnormal conditions in nuclear and process control plants, as part of probabilistic risk assessment studies (PRA).

HERAX approach to the solution of HRA problems, such as the identification of the likely operator errors and the assignment of probabilities to these errors, starts by asking what knowledge is used by a HRA experts in solving these tasks. This knowledge is then encoded in data structures and procedures that represent the knowledge explicitly, and that are separate from the inference procedures that apply this knowledge.

In essence, the system has a distributed architecture, being organized along the following modules or components:

1. A Knowledge Base Module.
2. An Inference Engine Module.
3. A Situation Data Base Module.
4. A User Interface Module:

An overview of HERAX's general architecture and the interaction between its components is shown in Figure 8.1.

The following subsections describe the functions associated with each of HERAX's modules.

8.7.1 Knowledge Base Module

HERAX's knowledge base consists of facts, rules and procedures. The facts used by HERAX to arrive at a conclusion come from two sources. The user inputs the basic information about the problem being considered including description of the operator tasks and the equipment used. These facts are compared with the facts included in the internal or resident knowledge base. Much of the acquired human reliability analysis knowledge, is stored and represented in HERAX's knowledge base module, as in any rule-based system, as rules or "production rules", that is, as conditional sentences relating statements of facts with one another.

To develop the knowledge base of the computer-based expert HRA system, the various human experts' problem-solving steps involved in carrying out the three existing HRA procedures, i.e., THERP, SLIM-MAUD, and APJ, have been modified, integrated, and then implemented into the computer. The major tasks of a HRA session that are performed by the computer system are described in subsequent sections.

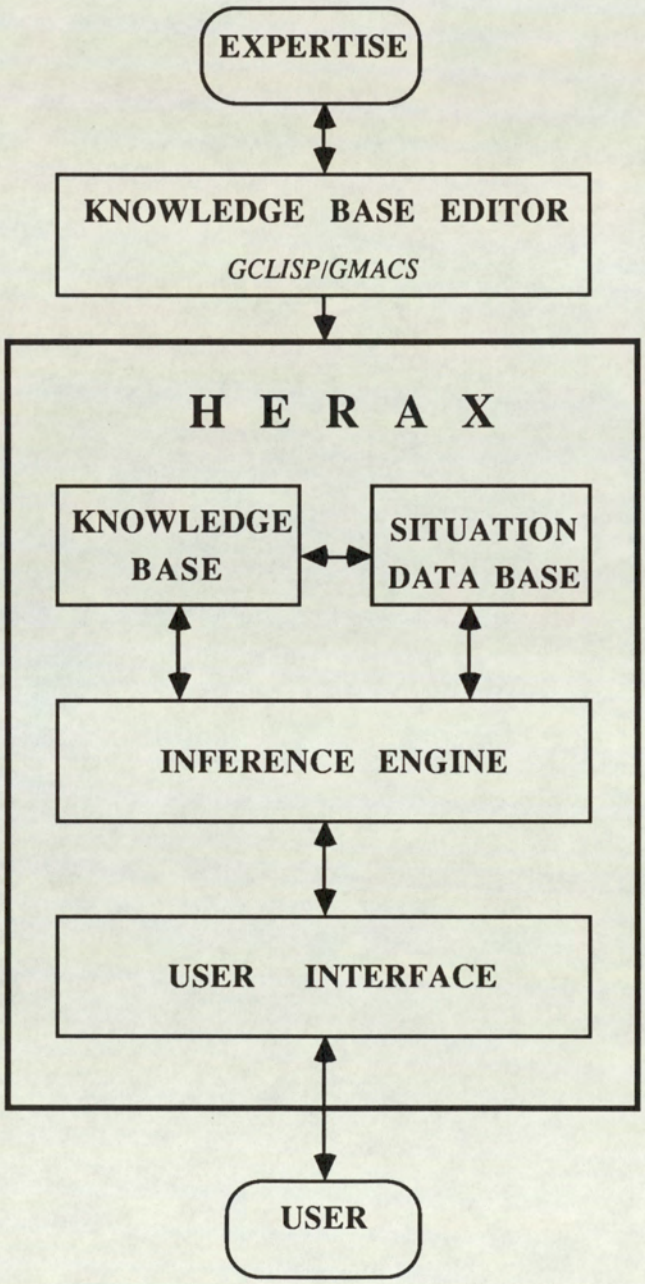


Figure 8.1 Overview of HERAX's Architecture

The knowledge base in HERAX has been designed to guide the analyst (the user) systematically through the HRA procedure according to the two main HRA stages identified in the literature review, i.e., human reliability modelling and quantification. While the knowledge represented is mainly of the three selected preventive techniques, the framework of the knowledge representation could be easily modified to integrate other analytical techniques.

The rules which constitute the knowledge base of HERAX are grouped into major categories or subparts of the HRA problem. These general rule sets or sub-knowledge bases are shown in Figure 8.2 and described in Table 8.1. The general classes of knowledge are then further subdivided into more specific rule sets or classes.

Each category of rules consists of a subset of rules which is generic for a specific HERAX step, e.g., rules for the identification of PSFs, rules for the identification of human error, and rules for the calculation of their likelihoods.

A major advantage of segregating the rules is that it enabled to achieve modularity, ease of maintenance and rapid access to the knowledge base.

The actions or conclusions parts of the rules are represented in terms of:

(a) ***Tables***, showing subtasks, major PSFs, and possible human error types.

(b) ***Event Trees***, used to represent the operator errors and their consequences (failure/success) on the system.

(c) ***Further questions***, used to ask the user for further input and description of the problem under study.

(d) ***Human Error Probabilities (HEPs)*** and their associated uncertainty bounds or error factors (EFs).

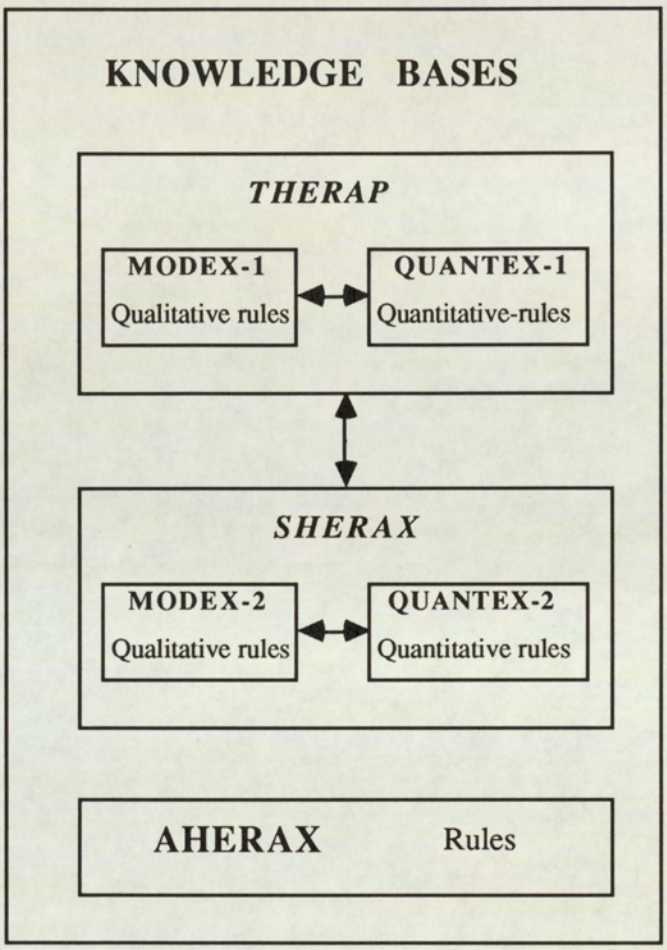


Figure 8.2 HERAX's Knowledge Base Module

Table 8.1 HERAX's Knowledge Base Categories

KNOWLEDGE BASE	APPLICATION
1. **THERAX-QUAL-RULES.** ***MODEX-1***	This knowledge base is applied to assist the user in identifying the possible human errors and their associated PSFs using THERAX.
2. **THERAX-QUANT-RULES.** ***QUANTEX-1***	This knowledge base is applied to support the user in quantifying human errors using THERAX.
3. **SHERAX-QUAL-RULES.** ***MODEX-2***	This knowledge base is applied to assist the user in defining and classifying the operator tasks, and identifying PSFs related to these tasks using SHERAX.
4. **SHERAX-QUANT-RULES.** ***QUANTEX-2***	This knowledge base is applied to assist the user in weighting, and rating the major PSFs, calculating the success likelihood indices (SLIs) for the tasks considered and then to converting these SLIs into probabilities.
4. **AHERAX RULES.**	This knowledge base is provided to assist the user in the application of the Absolute or Direct Probability Judgement method.

One example in English of the rules used to quantify the process operator action is shown in Figure 8.3.

RULE 75	
IF	1) the operator's required action is reading a display,
AND	2) the display is unannunciated, such as a meter, or a chart recorder,
THEN	1) the probability that the operator will fail to read the display is .003,
AND	2) the error factor (uncertainty bound) is 03.

Figure 8.3 Example of a HERAX's Rule in English Form

The translation of this rule into a LISP-like form is as shown in Figure 8.4.

RULE-75	
(**IF**	((Operator action = Read display)
AND	(Display type = Unannunciated))
THEN	((Conclusion = (HEP = .003)
AND	(EF = 03)))

Figure 8.4 Example of a HERAX's Rule in LISP-like Form

This is a simple rule in that it contains only two conditions and two conclusions. Some other rules are more complex, that is, composed of several conditions and several conclusions.

The rules in HERAX knowledge base are added, altered and deleted by the knowledge engineer through the ***"Knowledge Base Editor"*** using the GMACS text editor provided by the computer programming language GCLISP.

8.7.2 Inference Engine Module

The control mechanism or inference engine module in HERAX employs the information contained in the knowledge base to interpret the current contextual or task-specific data in the *working memory* or *situation data base* and is independent from the internal resident system knowledge base described in the previous section.

Rule Chaining in HERAX

Rule chaining refers to how the inference engine determines what rules are applicable based on what rules have already been applied. Thus the term *"chaining"* refers to the linking together of rules as the engine seeks to infer a solution. As discussed in Chapter 2, there are two forms of rule chaining - *forward chaining* (or *data-driven*) and *backward chaining*. HERAX uses a forward-chaining paradigm as its general search or control strategy operating in a *search-select-and-execute* fashion. Accordingly, the principal functions of the inference engine in HERAX are:

(i) Identification of applicable rules.

(ii) Conflict resolution.

(iii) Rule execution.

These functions are described below.

(i) Identification of Applicable Rules

The primary function of the inference engine is to find all rules that can be applied to the current problem at hand, i.e., to evaluate the premise of a rule, each condition at a time until they are exhausted. The premise of a rule consists of a set of conditions which are to be evaluated.

(ii) Conflict Resolution

The secondary function of the inference engine is to select from among the applicable rules a rule for application. More often than not more than one rule is applicable. In this case, a decision must be made as to which of these competing rules can be executed.

(iii) Rule Execution

The third function of the inference engine is to execute (fire) the action(s) or conclusion(s) specified by the selected rule. These actions might create new information (facts), which will be stored into the fact list in the situation data base, perform calculations, produce, as discussed above, some output (such as tables, event trees etc.), or ask the user for more input.

These three functions are continuously applied to the available rules until there are no more candidate rules to execute.

In summary, HERAX assumes nothing is known at the outset. To proceed, some data (facts) is required from the user. These facts are then stored into the fact list in the situation data base (working memory). Given this data, the inference engine can then find all those rules which are applicable. That is, those rules whose conditions are all

true. Some may execute creating additional information (facts) which will, in turn, cause more rules to be considered. This approach is used to impose a form of control over which sets of rules will be considered next by the inference engine. As an example, when analysing an operator tasks, modelling or identification of the operator errors should be performed before any type of quantification begins.

8.7.3 Situation Data Base Module

The situation data base or working memory, contains all the information (facts) which relates to the problem under analysis. This includes operator's tasks, type of analysis required, i.e., qualitative or quantitative, type of equipment used, and major factors which could influence the operator performance, such as experience, stress and use (or non use) of operating procedures.

8.7.4 User Interface Module

The user interface performs three main functions (see Figure 8.5). These are:

(a) Terminal control

(b) Fact verification

(c) Explanation generation.

(i) Terminal Control

The first main function of the user interface is the control of the computer terminal. The information displayed on the screen relates directly to the knowledge base. It permits the user either to enter specific commands or to select menu options. The commands are then fed to the inference engine that provides a mechanism for interpreting the commands and gaining access to the knowledge base.

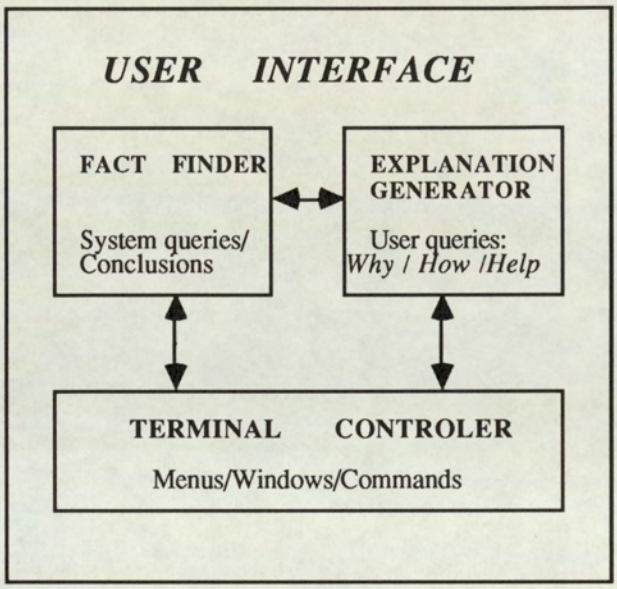


Figure 8.5 HERAX's User Interface

The user interacts with the system through windows/commands and menus. Two examples of menus are shown in Figures 8.6 and 8.7. The first menu, allows a user to select one of two activities to be performed by the expert system, either to start a HRA session with HERAX, or get an introductory explanation about the main computer program features. The third option is to exit the system. The user has the possibility to exit HERAX at any time during the analysis session.

(ii) Fact Verification

The user interface is also used as a fact-verifier which asks pertinent questions about the situation under consideration and displays the system's conclusions. These include:

information about the operator tasks/actions to be analysed, the type of man-machine interface (MMI), the design characteristics (ergonomics) of the equipment used, and the major factors likely to affect the operator performance.

The user interface utilizes three types of Prompt (or types of questions) procedures to acquire such information. These are described below.

HERAX
MAIN MENU
1. Introduction to HERAX
2. Start a session with HERAX
3. Exit to DOS
Any other command will display this list.
Please select a number and then press <ENTER>

Figure 8.6 HERAX's Main Menu

INTRODUCTION
1. Overview of HERAX 2. How HERAX works 3. Uses of HERAX use 4. Return to Main Menu
Please type the number of your choice:

Figure 8.7 HERAX's Introduction Menu

The Prompt1 procedure poses a question and expects only a '*Yes/No*' response from the user (an example is shown in Figure 8.8).

Identification of Tasks	
Task1 =? Task2 =? ... More tasks =?	
Yes/No	Ctrl-Break to exit

Figure 8.8 An example of Yes-No questions

The Prompt2 procedure expects a single-valued response from the user. In order to assist the user to respond effectively, a list of multiple-choice questions is also presented. All the questions are numbered and the user is expected to enter the number of the question selected (an example is shown in Figure 8.9).

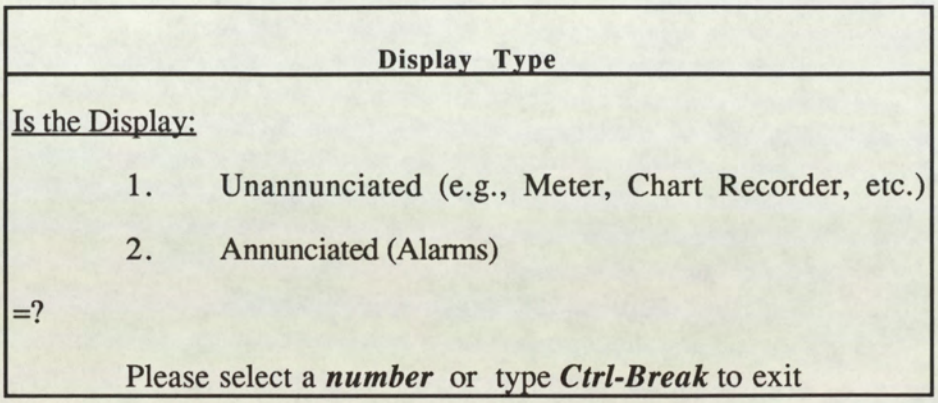


Figure 8.9 An example of multiple-choice questions

The Prompt3 procedure displays a question and expects a numerical value from the user, as when rating the effect and weighting the importance of PSFs on a scale that ranges from 0 to 100 (an example is shown in Figure 8.10).

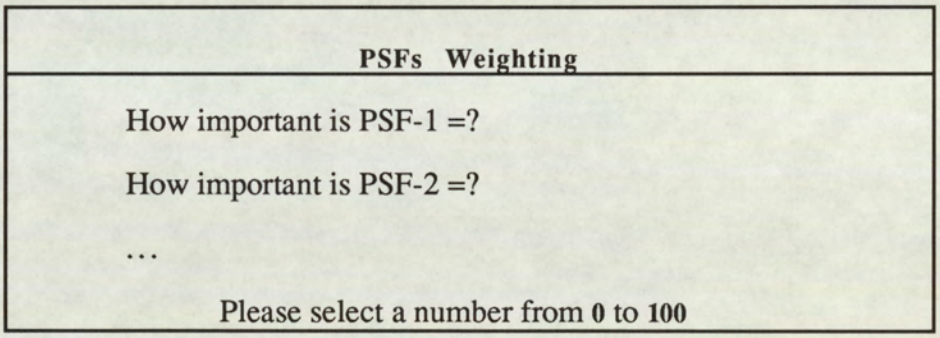


Figure 8.10 An example of numerical-value questions

(ii) Explanation-Generation

A primary requirement for an acceptable expert system is that it should have a facility to explain its lines of reasoning. Representation of knowledge in HERAX using rule-based technique has simplified the implementation of such facility.

The explanation system offers three options to the user of HERAX, and each of these is described below.

1. Explanation of ***What*** and ***How*** HERAX Works

Before starting a session with HERAX, the user is provided with the possibility to get an introductory explanation about the structure of the system, how it works, what are its capabilities and uses, and finally, what type of data is required to use the system and where to get it.

2. Explanation of ***Why*** certain facts are required

The Explanation-generation module is able to explain to both the expert and the user *why* a specific question was asked. This facility allows users to see the reasonings which are hidden behind a question and it is a useful feature in cases where the users are unaware of the likely implications which may be caused by a specific response. A prime advantage of such a facility is that a user can have a wider understanding of the logics and dependencies before making any specific commitment. A user may ask *"Why"* certain information is needed and the Explanation-generation system will access the knowledge base to retrieve and display all the relevant rules which are directly related to a particular question.

3. Explanation of ***How*** certain facts are deduced

The explanation-generation module is also able to explain the reasoning, i.e., displaying the rules, that led to a specific conclusion (or to a further question by the system). This ability of HERAX to answer the user's *"Why"* and *"How"* questions is important, for it increases the user confidence in the system's decision-making ability.

4. *Help* Facility

The user is also provided with a facility in form of a detailed information in answering some of the system questions.

8.8 HERAX Approach

The HERAX approach to the modelling and quantification of human reliability as part of PRA studies uses three main procedures:

1. THERAX.- *"Technique for Human Error/Reliability Analysis eXpert"*
2. SHERAX.- *"Subjective Human Error/Reliability Analysis eXpert".*
3. AHERAX.- *"Absolute Human Error/Reliability Analysis eXpert".*

Figure 8.11 shows the general flow of decision in HERAX procedure. Each of the THERAX, SHERAX and AHERAX steps is described in more detail in the following subsections.

The HERAX analytical process starts by guiding the user through the collection of specific plant data and the identification of potentially important human tasks and PSFs. Then to reduce (screening) the number of human tasks and select the key ones for more detailed analysis.

Having collected the necessary data, the user is presented with the THERAX technique which includes a human error probabilities (HEPs) data bank.

In case not all the HEPs have been generated using the THERAX procedure, the expert system recommends the use of other quantification techniques to generate the missing data. HERAX then proposes two subjective techniques, namely SHERAX and

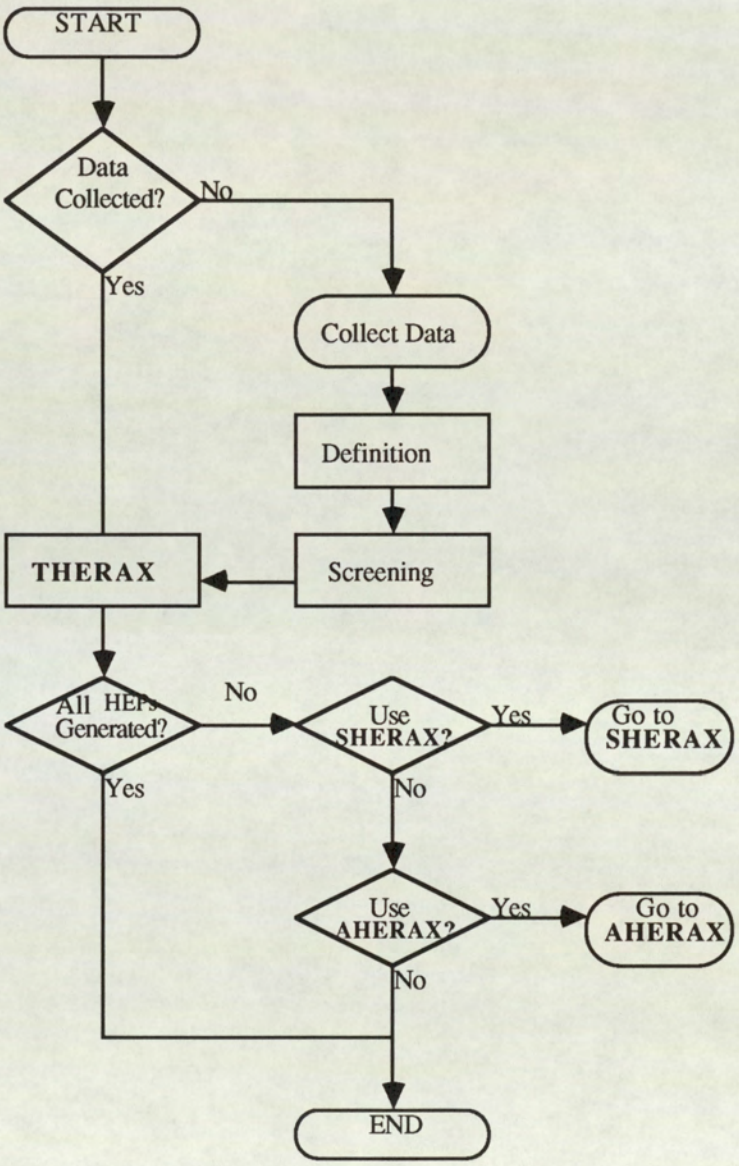


Figure 8.11 General HRA procedure using HERAX

AHERAX, and guides the user in the selection and application of the most appropriate one to the case under study.

8.8.1 THERAX Procedure

THERAX is a systematic and interactive method based on the main tasks an expert in human reliability analysis would perform when using THERP technique.

The decision process involved in the application of the THERAX procedure is illustrated in Figure 8.12.

The main uses of the computer program THERAX can be summarized as follows:

1. to provide guidance to the user in the collection of plant data and the identification of the operator tasks and PSFs.
2. to analyse the operator tasks and predict the possible human errors;
3. to provide a list of the major PSFs;
4. to retrieve and assign probabilities (HEPs) to these errors;
5. to provide explanation and justification of the above.

THERAX procedure uses the two following models, which correspond to the two main stages of the HRA procedure, i.e., modelling and quantification of human reliability: -

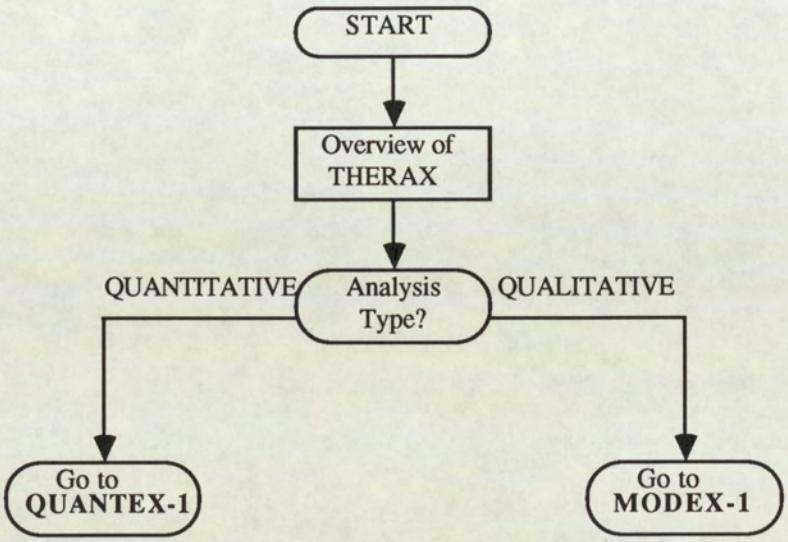


Figure 8.12 Decision Flow in THERAX

1. MODEX-1 - This program is used to model human reliability.

2.QUANTEX-1 - This program is used to quantify human reliability, and it incorporates a data bank on human error probabilities (from the Handbook by Swain et al, 1983).

A. Human Reliability Modelling with *MODEX-1*

The main steps involved in MODEX-1, the human errors modelling procedure, as part of HERAX, using the THERAX technique are shown in Figure 8.13, and are described in more detail below.

Step 1 - Task Analysis. This MODEX-1 step involves the breakdown of complex tasks into subtasks/actions in association with the type of Man-Machine Interface (M-MI) used.

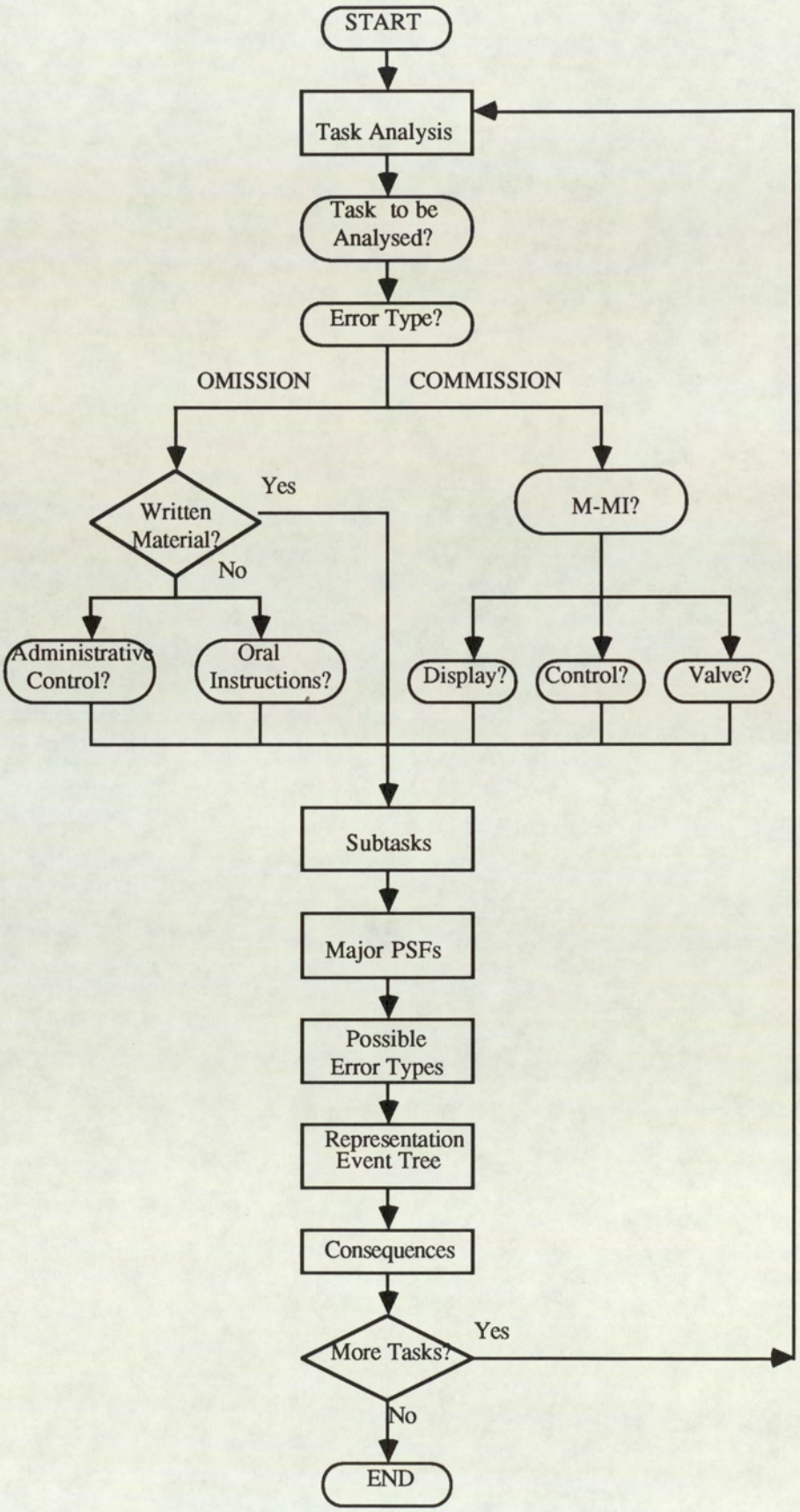


Figure 8.13 Steps and decision flow in MODEX-1

Step 2 - Task Definition. MODEX-1 requests the user to describe the task to be analysed.

Step 3 - Type of Errors. The user is requested to select the type of errors, either *Omission* or *Commission*, for MODEX-1 to analyse. If the user selects errors of Commission, then the next question the system will ask is determination of the type of *M-MI*, i.e., Displays, Controls, or Valves.

If the user selects errors of Omission, then MODEX-1 asks whether performance of the task of interest requires the use of *Written Procedures* . If the answer is *Yes*, then the system displays the conclusion it has reached. If the answers is *No*, then the user is asked to determine whether *Administrative Control* or *Oral Instructions* are being used.

The above user information will enable MODEX-1 to draw its conclusions as described below.

Step 4 - Subtasks. The system identifies the main subtasks/steps of the task being analysed in relation to the type of M-MI used.

Step 5 - Performance Shaping Factors. MODEX-1 identifies and provides the user with a list of the major PSFs associated with each subtask/M-MI.

Step 6 - Human Errors Identification. MODEX-1 identifies the possible (omission or commission) types of human errors for each task step/M-MI, such as *reading* and *selection* errors.

Step 7 - Human Error Representation. Logical representation of the human actions/ errors identified for subsequent quantification using Event Trees.

Step 8 - Human Error Impacts/Consequences Assessment. Modex-1 determines the consequences of the identified errors.

A. Human Reliability Quantification with *QUANTEX-1*

The main steps involved in QUANTEX-1, the human reliability quantification expert, as part of HERAX, using the THERAX method are shown in Figure 8.14, and are described in more detail below.

Step 1 - Situation Definition. QUANTEX-1 starts by asking the user to determine whether the situation under study is *Normal* (routine operations, such as maintenance and testing) or *Abnormal.* (start up, shut down, emergency procedures, etc.).

Step 2 - Task Analysis. If the situation of interest is Normal, the analysis starts by modelling the human operator performance using the MODEX-1 procedure as described above. Then, a detailed task analysis is carried out to describe further the operator actions and design of the equipment used.

Step 3 - Actions Classification. If the situation is Abnormal, then the user should determine whether performance of the task considered involves diagnosis or rule based actions. If diagnosis is involved, QUANTEX-1 prompts the user to indicate the time elapsed before the operator takes action in response to the first alarm. If however, rule based actions are involved, the steps in MODEX-1 are applied followed by a detailed task analysis.

Step 4 - Human Error Probabilities. QUANTEX-1 generates the required nominal HEPs and their error factors (EFs) or uncertainty bounds for the operator actions identified in steps 2 and 3 above.

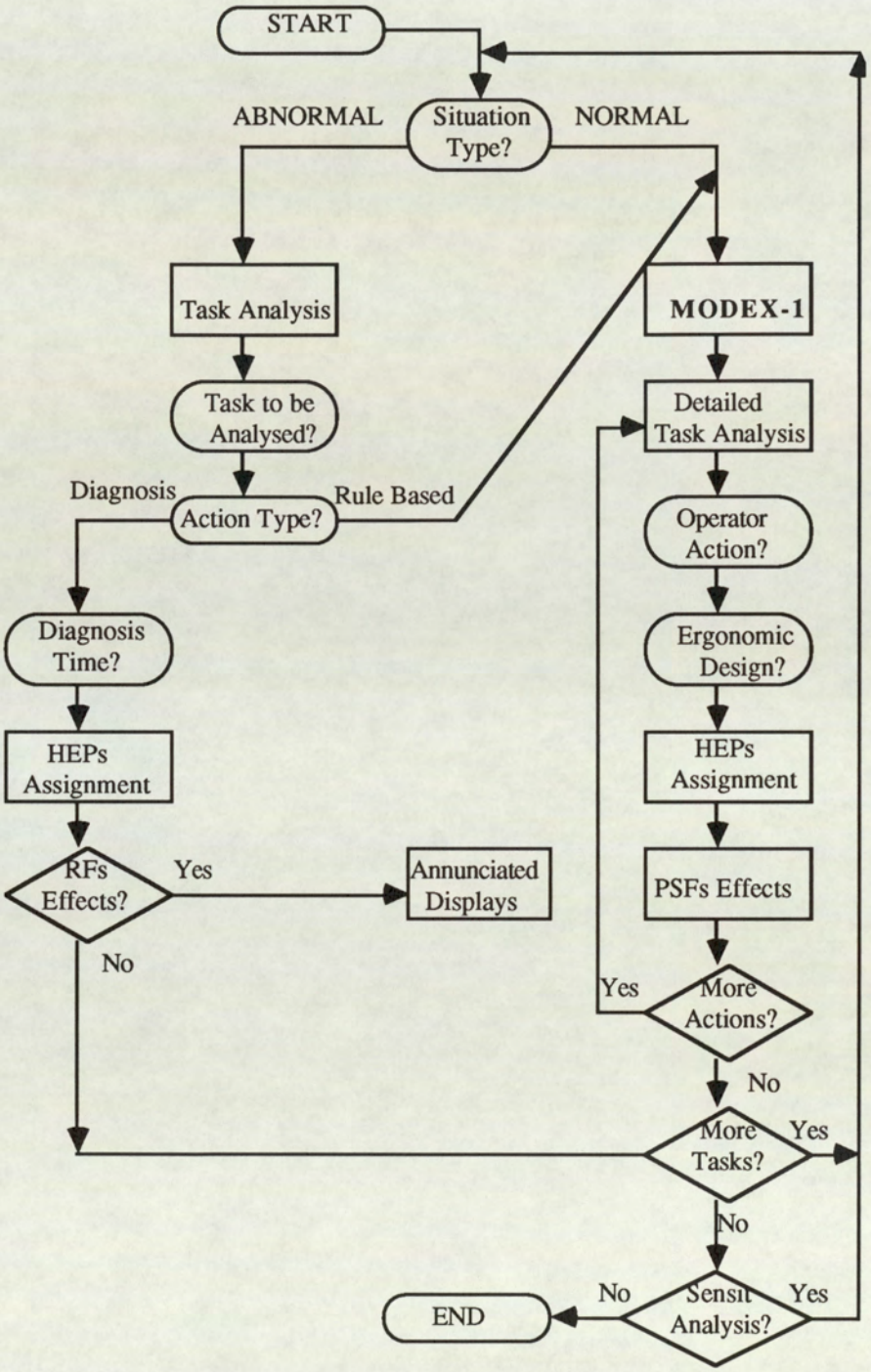


Figure 8.14 Steps and Decision Flow in QUANTEX-1

Step 5 - Effects of PSFs. The two major effects of PSFs on the original HEPs assessed by QUANTEX-1 are stress and experience.

Step 6 - Effects of Dependences. Quantex-1 calculates the effects of dependences between tasks as well as between operators on the error probabilities.

Step 7 - Effects of Recovery Factors. Assessment of the effects of RFs (Annunciation of deviant conditions) on the original HEPs.

Step 8 - Total Probability of Failure. QUANTEX-1 uses Event Trees to calculate the total probability of failure, taking into account the effects of PSFs, dependences, and RFs.

Step 9 - Sensitivity Analysis. If sensitivity analysis is required, the user re-starts the THERAX procedure from the beginning in order to make the necessary modification to the assumptions made during the previous analysis session.

8.8.2 SHERAX Procedure

SHERAX is a modified version of the SLIM-Maud technique (Embrey et al., 1984). In cases where not all the required HEPs for the tasks considered have been generated by using the THERAX method, the user of the HERAX analysis system is guided in the selection and use of either the systematic subjective method SHERAX, or the absolute probability judgement AHERAX.

The decision process involved in the SHERAX procedure is illustrated in Figure 8.15.

SHERAX procedure is used within HERAX mainly to perform the following tasks:

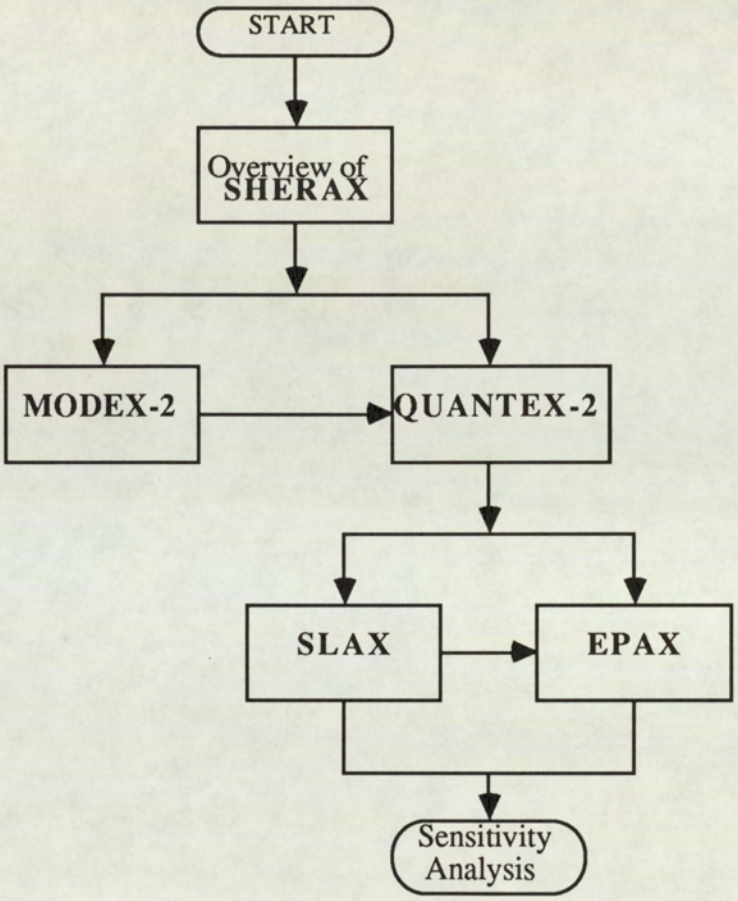


Figure 8.15 Decision Flow in SHERAX

1. to provide guidance to the user in the definition of situation and identification of operator tasks and PSFs;
2. to assist the user in the rating and weighting of PSFs,
3. to calculate the success likelihood indices (SLIs) for all the tasks identified,
4. to calculate the error probabilities for all the tasks considered;
5. to provide explanation and justification of the above.

The main steps involved in the SHERAX procedure can be classified, similarly to THERAX, into two main stages: modelling and quantification . These two functions are performed by two models:

1. MODEX-2 - This model is used for the modelling of human reliability.

2. QUANTEX-2 - This model is used to quantify human reliability, i.e., to calculate the success likelihood indices and error probabilities.

A. Human Reliability Modelling with *MODEX-2*

The main steps involved in MODEX-2, the human errors modelling procedure, as part of HERAX, using the SHERAX technique are shown in Figure 8.16, and are described in more detail below.

Step 1 - Identification of tasks. The user identifies major PSFs and the operator tasks for which HEPs are needed.

Step 2 - Classification of Tasks. The system provides guidance to the user in the grouping of the identified tasks into three categories: "Skill-based", "Rule-based", or "Knowledge-based".

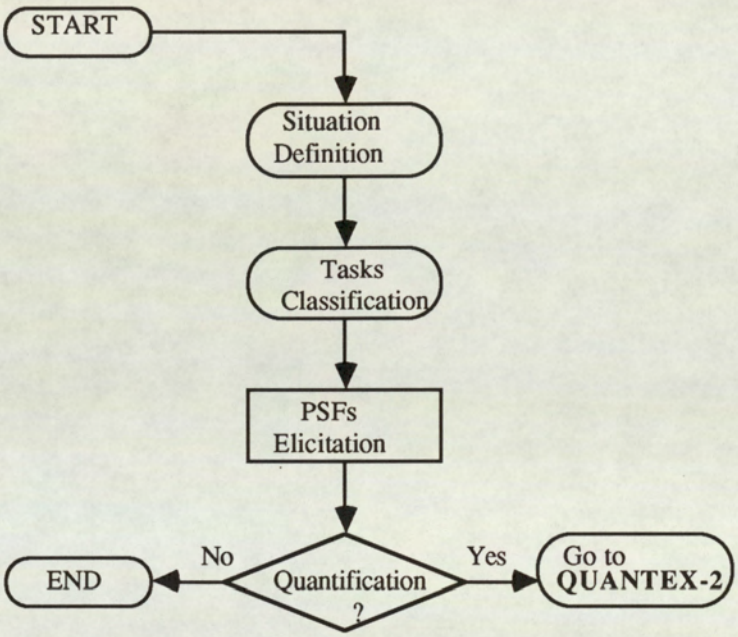


Figure 8.16 Steps and Decision Flow in MODEX-2

Step 3 - Identification of PSFs. Users are provided with a list of PSFs from which to select the most relevant ones, however they are advised to identify those PSFs which are not on the list but which are important in the current situation.

B. Human Reliability Quantification with *QUANTEX-2*

QUANTEX-2 is comprised of two sub-modules:

1. **SLAX**: **S**uccess **L**ikelihood **A**nalysis e**X**pert
2. **EPAX**: E**r**ror **P**robability **A**nalysis e**X**pert

The main decision tasks involved in QUANTEX-2, the human reliability quantification expert, as part of HERAX, using the SHERAX technique are shown in Figure 8.17:

A. Success Likelihood Index Calculation using **SLAX**

Step 1 - Rating of the effects of PSFs. SLAX asks the user to rate the identified PSFs, according to their effect on the operator performance of the task considered, on a scale provided by the computer program that ranges from 0 (for lowest effect) to 100 (for highest effect).

Step 2 - Weighting of the importance of PSFs. As in the previous step, the user is requested to weight and rank the specific PSFs according to their importance for the operator performance on a scale provided by the program that ranges from 0 (for least important) to 100 (for most important).

Step 3 - Calculation of Success Likelihood Indices (SLIs). This task is performed by SLAX program based on the user's ratings and weightings of the relevant PSFs.

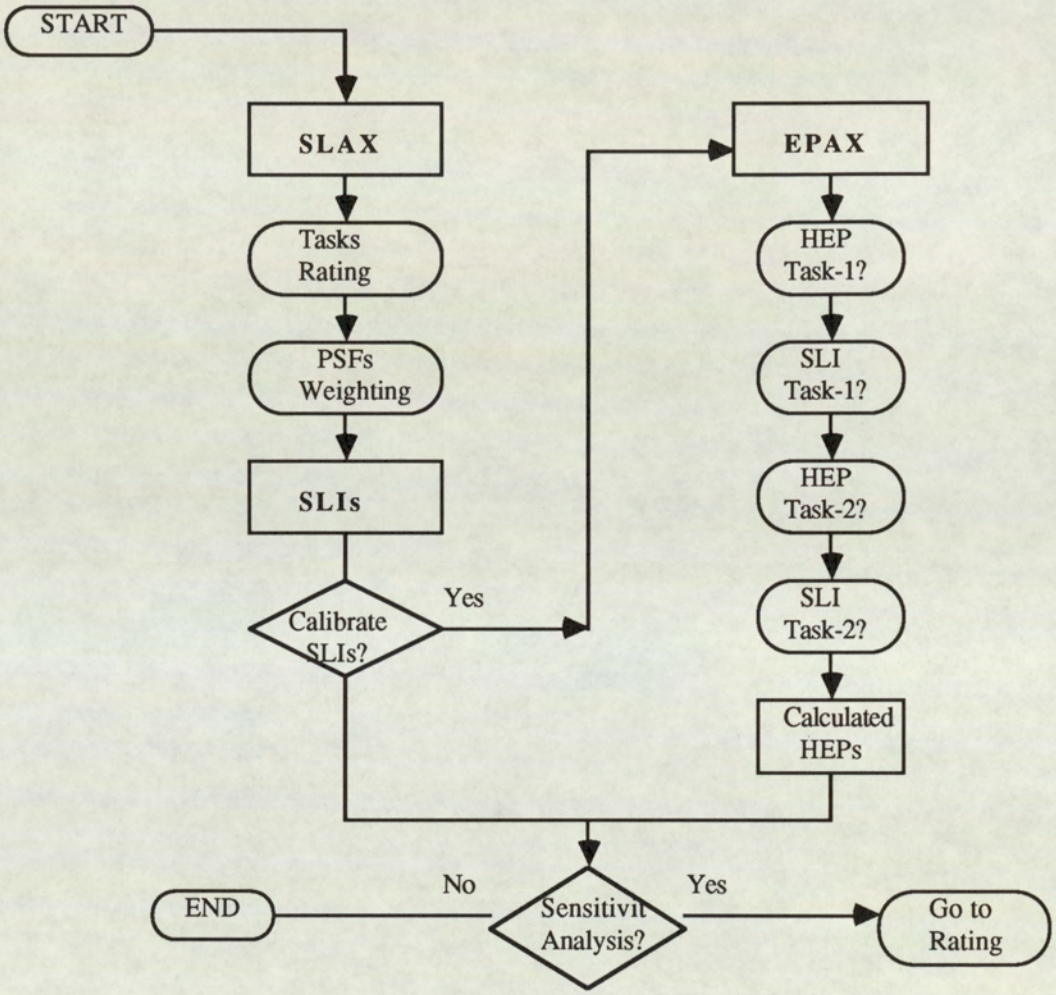


Figure 8.17 Steps and Decision Flow in QUANTEX-2

Step 4 - Conversion of SLIs into HEPs. The user is requested to supply two known HEPs and SLIs, EPAX then deduces the required probabilities for the remaining tasks.

Step 5 - Sensitivity Analysis. If sensitivity is warranted the user re-starts SHERAX.

8.8.3 AHERAX Procedure

This technique has been added to the expert system only recently, instead of SHARP technique which was originally implemented in the expert system in combination with THERP. Not all steps have been implemented yet.

AHERAX.procedure is based on the APJ technique described in Chapter 7. It can be used either by a single expert, or group of experts. However, the user is advised not to rely upon a single expert judgements, but instead, use a group of experts to avoid biases. . The overall procedure is shown in Figure 8.18. The main steps are as follows:

1. Selection of subject matter experts
2. Preparation of task statements
3. Preparation of response booklets
4. Development of instructions for subjects
5. Obtention of judgements
6. Aggregation of individual estimates

The above AHERAX steps are described in more detail below.

Step 1 - Selection of Subject Matter Experts. The user is advised to select experts for making judgements that are familiar with the plant/system and tasks of interest. An example is presented.

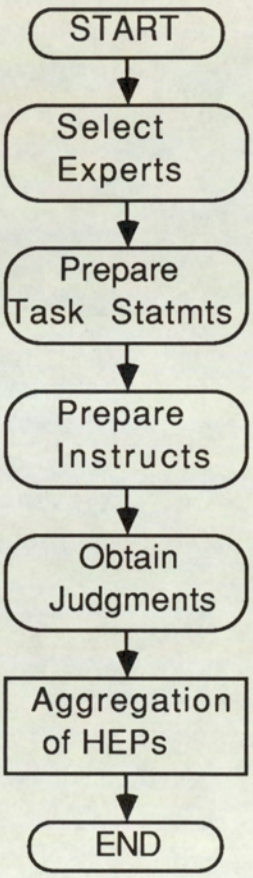


Table 8.18 Steps and Decision Flow in AHERAX

Step 2 - Preparation of Task Statements. The system stresses the critical aspect of well-defined task statements and assumptions (PSFs) to this procedure. It then advise the user on how to determine the level of detail.of tasks.

Step 3 - Preparation of Response Booklets. An example of scale on which the experts will indicate their judgements is suggested to the user. Guidelines about the contents and layout of the booklets are also given.

Step 4 - Development of Instructions for Experts. The user is provided with guidelines about what the instructions given to experts at the beginning of a session should indicate.

Step 5 - Obtention of Judgements. The user is asked to collect estimates from the experts. If a consensus group is being utilised, the user is advised to use a facilitator to overcome any personality/group problems and to prevent any biases from affecting the experts judgements. Calculation of inter-judges consistency is not currently implemented in AHERAX.

Step 6 - Aggregation of Individual Estimates. The system asks the user to enter the individual estimates obtained from (not more than four judges) for the (not more than four) tasks of interest, then the computer system calculates the statistical aggregation of the each individual probability. Upper and Lower Uncertainty Bounds (UBs) estimations are not yet implemented in AHERAX.

8.9 CONCLUSION

This chapter has first discussed the reasons for, and the benefits to be gained from, developing an expert system approach for human reliability analysis. The main conclusion that can be drawn from this discussion is that there are a number of practical reasons why the expert system approach may be particularly advantageous to HRA. In particular, HRA is a scarce, complex, expensive, yet critical component of any systematic safety/reliability or risk assessment study. Therefore, there is a need for a tool that would overcome some of these problems, i.e., make the knowledge and expertise of HRA practitioners more widely available and easy to use by non-experts. This objective could be well achieved by using AI and expert system approaches.

The chapter has also described the main features of HERAX, a rule-based expert system approach for human reliability analysis. The HRA expert system approach described here uses Artificial Intelligence (AI) and expert system techniques similar to some extent to those used in most of the other expert systems that have been built in the related areas of health, safety and reliability (see Chapter 2), such as separation of the knowledge base from the control procedures that use that knowledge, explanation capabilities, and application of the system development cycle procedure. However, the proposed expert system is distinctive from the previous expert systems in many respects. Some of the major characteristics of HERAX are:

- One important distinctive feature of the proposed expert system is that it can be used for the modelling and quantification of process operator tasks under normal as well as abnormal process control plants conditions.

- A second major distinctive feature of this expert system is that it incorporates three existing HRA procedures. These three techniques were altered so that they could

be used by users non-experts in ergonomics/human factors and human reliability analysis. In that respect, the expert system developed herein serves the important function of, and provides a framework for, integrating existing HRA methods into one systematic, interactive and easy to use approach.

- A third, and perhaps most important, feature of this system is the incorporation of a formal data bank on human error probabilities and a list of major performance shaping factors (PSFs).

- A fourth distinctive feature of this system is that it has been implemented on a micro-personal computer. This will increase its portability and make its dissemination among non-experts in ergonomics/human factors, as well as computing, easier and inexpensive.

- Finally, this expert system includes a facility that it is capable of providing various forms of explanations to the user about its line of reasoning. Other features of the system are described in the previous sections.

Some of the expected benefits of the use of HERAX in the field of HRA can be summarized as follows:

- To simplify the performance of task analysis and human reliability analytical procedures.

- To reduce the time and cost for analysis.

- To improve human reliability analysis consistency and accuracy by following a structured, logic flow each time a given problem is considered.

- As an educational tool, its explanation facilities are a valuable assest as they allow users to appreciate precisely how experienced people are able to reach a solution more rapidly, by asking the right questions in the right order. By learning such techniques, inexperienced people are able to understand how and why certain rules of thumb have been taken.

- The guidance facilities used by HERAX for extracting and manipulating information required can prevent many unrealistic analysis, thereby increasing the speed of arriving at a suitable result. The tasks of the analyst have been made much easier using these procedures. It further avoids the difficulties which are associated with the selection and application of HRA methods.

- To enhance the utility of existing HRA methods as well as provide an effective means to encourage systematic analysis of human reliability. Thus HERAX could contribute to industrial safety by providing a consistent and integrated framework for enhancing safety in design and operation.

- Finally, although HERAX's domain is confined to the analysis of process plants operators reliability, the system can be easily modified or adapted to other situations and environments. This capability is facilitated by the use of production rules as a knowledge base representation.

The next chapter presents a case study analysis that demonstrates the use of HERAX to the analysis of typical abnormal situations in chemical process plants.

CHAPTER 9

APPLICATION OF HERAX

A CASE STUDY: RUNAWAY REACTION

9.1 Introduction

Many of the early hazard/risk analyses ignored the operators, assuming that the operators would always carry out what is required of them. Other analysts went to the other extreme, assuming the operators would always fail. The current tendency among process managers and designers is to *'design out'* the human element in process control and rely more on protective or fully automatic systems. It is believed that the use of automatic protective systems should reduce the possibility of human errors and hence ensure a higher level of plant safety and productivity. The positive aspects of human performance in monitoring and mitigating failure consequences have been under-rated by plant designers.

Although automation is being extensively applied, recent major accidents in the aviation and process industries show that there arise situations where the operator's skills such as diagnosis, decision making and manual intervention are required, particularly in the cases of unforeseen problems, e.g. start-up, shut-down and emergency procedures.

There is, therefore, a need for tools to assist plant management and designers to:

(i) evaluate more accurately the impact of human reliability on plant safety and reliability.

(ii) assess the safety integrity of automatic protective systems and human reliability in order to determine the optimum balance between automatic systems and human performance.

In this chapter a case study, *"runaway reaction"*, is presented to demonstrate the application of **THERAX** procedure, one of the three procedures included in the expert system HERAX, to the identification and quantification of human reliability in abnormal process situations, and to examine the effects of reliance upon automatic protective systems as means for increasing process safety.

The ideal, of course, is to use an actual case. But, because of the many unsuccessful attempts made to get a financial support from industrial companies (such as British Gas and BIP Chemicals) for the practical application of the system on real life situations, and because of time limits, the work carried here involves a study of a hypothetical case example.

9.2 What is a Runaway Reaction?

Runaway reaction describes the situation when for any reason an exothermic system becomes uncontrollable. This is often linked with coolant failure on a reactor vessel, but can occur from self-heating at any stage of a process. A runaway reaction leads to a rapid increase in temperature an pressure which can rupture the containing vessel. The runaway happens because the rate of reaction, and hence the rate of heat generation, increases exponentially with temperature, whereas the rate of cooling increases only lineary with temperature. Once heat generation exceeds available cooling the rate of temperature rise becomes progressively faster (Tharmalingam, 1989).

Incidents involving exothermic reactions continue to cause problems when the reaction mass overheats in an uncontrolled manner and overpressurisation and loss of containment result. Runaway reactions are often the cause of major industrial accidents; Bhopal was an extreme example. In the UK in 1986, 31 such incidents were reported resulting in injury to 18 individuals (Pantony et al., 1989). According to Tharmalingam (1989), there has been an average of ten reported incidents per year in the UK alone over the last 20 years.

9.3 Original Process Design: *Human Control*

The proposed plant design (shown in Figure 9.1) is intended to carry out an exothermic reaction between chlorine and a hydrocarbon. A runaway reaction will occur if the chlorine flow is too high or the hydrocarbon flow is too low.

The reaction normally occurs at 360°K. A high temperature alarm is used to warn the operator if the temperature rises to 375°K. The operator should respond as quickly as possible by closing either valves V-1 or V-2 on the chlorine feed line. The time available to act depends on the rate of reaction. If the temperature reaches 395°K or above, then a runaway reaction is expected to occur.

While the chlorine feed is automatically controlled, the design depends on the process operator to monitor the temperature indicator (TI) and to respond to the high temperature alarm (TA).

9.3.1 Objectives of the Analysis

The main objective of the analysis is to determine the effect of the operator failures on the overall plant safety/reliability, i.e., to determine the following:

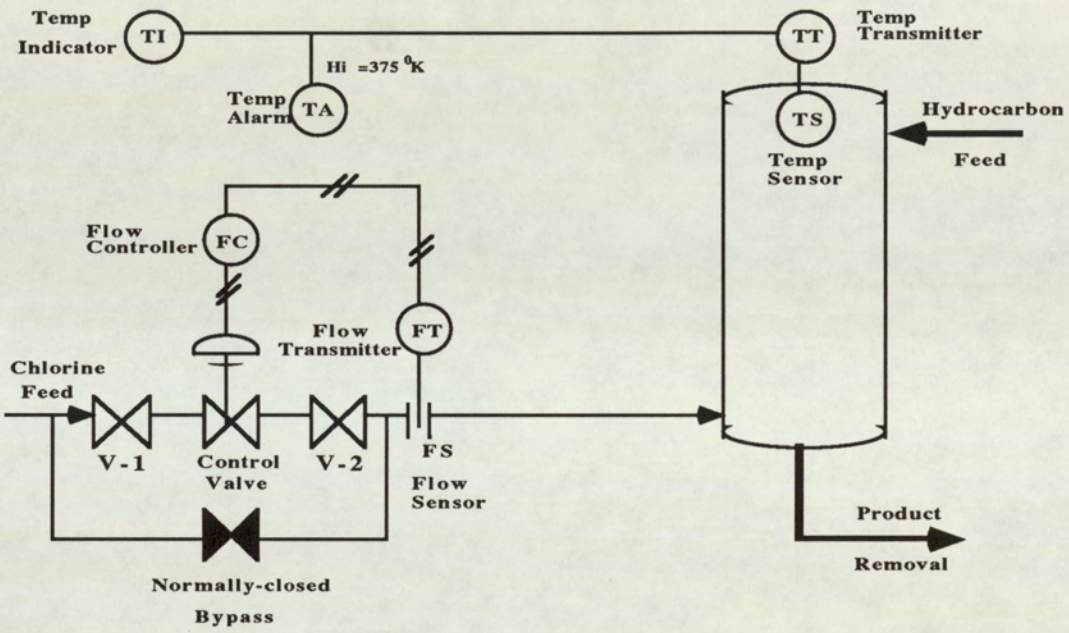


Figure 9.1. Original plant design.

1. What is the probability that the operator does not *Detect abnormal rates* of temperature rise, or temperature reaching 375^{o}K ?

2. What is the probability that the operator fails to *Stop the chlorine flow* before a runaway reaction occurs ?

9.3.2 Assumptions

The following assumptions have been made:

1. The process operator is required to continuously monitor the changes in temperature.

2. A high temperature of (375^{o}K) is assumed to have been initiated by high chlorine flow.

3. If the temperature rises too rapidly or reaches 375^{o}K, the operator is instructed to use the plant operating procedures and initiate *Immediate Actions* for the prevention of a runaway reaction accident; therefore, the primary interest in this case is responding to the alarm. No diagnosis is involved.

4. The plant operating procedures stipulate that whenever the chlorine flow is too high and the temperature reaches 375^{o}K, the operator should (in addition to other tasks not considered here) stop the chlorine flow by going out of the control room as quickly as possible and turning off either valves V-1 or V-2 manually.

5. The operating procedures instructions must be performed within 10 minutes after high temperature alarm.

6. It is assumed that two operators are involved in the emergency procedure.

7. The two operators are assumed to be skilled, and operating under optimum stress level.

8. All the Human Error Probabilities (HEPs) are taken from the Handbook by Swain et al. (1983).

9.3.3 Tasks Identification

To avoid a process runaway reaction, the following main operators tasks have been identified.

Operator (A) to:

1. Perform monitoring of reaction temperature changes.
2. Detect abnormal rates of rise in temperature or high temperatures.
3. Verify (check-read) Temperature Indicator (TI) after corrective measures are taken by operator (B).

Operator (B) to:

1. Refer to the appropriate written procedures to cope with abnormal events and carry out the activities indicated.
2. Stop the chlorine flow by turning-off either valves V-1 or V-2 in order to prevent a runaway reaction.

9.3.4 Expert System Analysis

The following shows the interaction between the expert system and the user. Note that *bold* characters are the expert system's queries, followed by the user's input, whereas *boxes*, *tables*, and *event trees* are the conclusions the system has reached. The author's comments are between *parentheses*.

****

Have you Collected the necessary plant and operator data? Yes

THERAX: Technique for Human Error/Reliability Analysis eXpert

Type of Analysis?	Quantitative
Type of Situation?	Abnormal
Task to be analysed?	Monitor
Type of Actions?	Rule-Based
Type of Errors to be analysed?	Omission
Administrative Control used?	Yes
Operator Action?	Initiate a scheduled checking/inspection

Task being considered:	Monitor	
OMIT INITIATE ACTION	HEP = 0.001	EF=3

****

More Tasks?	Yes, Detect Rising Temperature	
Type of Errors?	Commission	
Man-Machine Interface?	Displays	; (A Table, and an Event Tree, showing Subtasks/PSFs and Errors, are displayed)
Type of Displays?	Annunciated	; (Alarms)

Table1 Use of ANN Displays

SUBTASKS	MAJOR PSFs	ERROR TYPES
1. RESPOND: Initiate Action	- Number of alarms	- Omission: no Act
2. SCAN	- Number of unimportant indicators	- Timely response error
3. READ Message	- Number of false alarms	- Scanning error
4. DIAGNOSE	- Design of indicators	- Reading error
	- Stress	

; (an Event Tree is also displayed)

****

Operator Action?	Respond
Number of ANNs?	7

Task being considered:	Detect Rising Temperature	
FAIL NOTICE ONE of 7 ANNs	HEP = 0.009	EF=10

****

More Actions? No

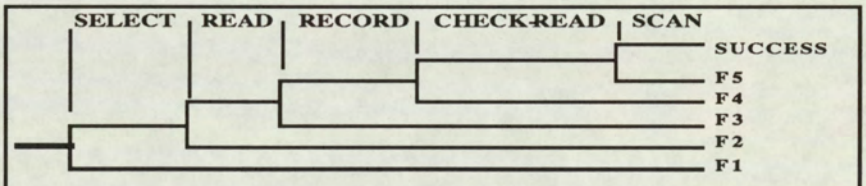
More Tasks? Yes; Detect Rising Temperature

; (same as above, however the operator having been alerted, will now look at the TI)

M-MI? UNANnunciated Displays

Table2 Use of UNAN. Displays

SUBTASKS	MAJOR PSFs	ERROR TYPES
1. SELECT.	- Stress.	- Selection error.
2. READ.	- Signal-processing rate.	- Reading error .
3. RECORD.	- Frequency of scan.	- Recording error
4. CHECK-READ.	- Use of written material.	- Check-reading error
5. SCAN.	- Relationship to ANNs.	- Scanning error .
	- Content of information.	
	- Ergonomics of displays	



HRA Event Tree 1

****

Operator Action? Check-Read

Type of UNAN? Digital Indicator

Task being considered:	Detect Rising Temperature	
FAIL CHECK-READ	HEP = 0.001	EF=3

****

More Actions? No

Task being considered:	Detect Rising Temperature	
1. FAIL NOTICE ANN	HEP= 0.009	EF=10
2. FAIL CHECK-READ	HEP= 0.001	EF=3
TOTAL PROBABILITY =0.010		

****

More Tasks? Yes, Stop Chlorine Flow

Type of Errors? Omission (Omit Initiate Action)

Written Materials Used? Yes

Table3 Use of Written Material

SUBTASKS	MAJOR PSFs	ERROR TYPES
1. PREPARE (write) procedures. 2. USE Procedures. 3. PERFORM Steps.	- Quality of procedures. - Type of procedures. - Dependence among steps. - Check-off provisions. - Number of steps. - Stress.	- Omit Step. - Incorrect writing. - Omit use procedure.

(An Event Tree will also be displayed showing the Subtasks and Consequences of errors)

****

Operator Action? Use Procedural Items

Task being considered:	Stop Chlorine Flow	
OMIT ITEM	HEP = 0.001	EF= 3

****

More Tasks? Yes, Stop Chlorine Flow

; (Same as above, but this time errors type is commission).

Type of Errors? Commission

Man-Machine Interface? Manual Valve

Table4 Use of valves

SUB-TASKS	MAJOR PSF's	ERROR TYPES
1. SELECT Valve. 2. MANIPULATE (Open/Close) 3. DETECT (valve stuck)	- Written procedures - Number of valves. - Location - Adjacement - Quality of labeling. - Identification of valves - Configuration of valves. - Status indicators - Stress.	- Omission - Selection Error. - Reversal Error - Detection Error

****

Operator Action? Select (Right valve)

Valve Design? Unclearly labeled, part of a group of 2 or more valves similar in ALL the following: size, shape, state, and presence of tags.

Task being considered:	Stop Chlorine Flow	
FAIL SELECT	HEP = 0.01	EF=3

****

More Actions? Yes, Manipulate (Close valve)

Task being considered:	Stop Chlorine Flow
REVERSAL ERROR	HEP = 0.0001

****

More Actions? No

Task being considered:	Stop Chlorine Flow	
FAIL INITIATE ACTION	HEP= 0.001	EF= 3
FAIL SELECT VALVE	HEP= 0.01	EF=3
REVERSAL ERROR	HEP= 0.0001	EF=3
TOTAL PROBABILITY = 0.0111		

****

Assessment of PSFs Effect? Yes

Stress Level? Moderately High

Experience Level? Skilled

STOP CHLORINE FLOW	HEP= 0.022	EF=3

; (the HEP was multiplied by a factor of 2)

****

More Tasks? Yes, Verify ; (Temperature back to normal)

Type of Error? Commission

M-MI? UNAN Display ; (Table3 and an Event tree are displayed)

Type of UNAN? Digital Indicator

Operator Action? Check-Read

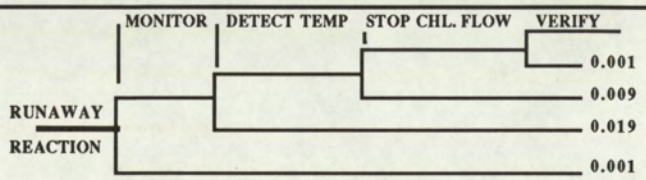
Task being considered:	Verify Temperature	
FAIL CHECK-READ	HEP= 0.001	EF=3

****

More Tasks? No

Tasks being considered

1. Perform Monitoring
2. Detect rate of rise in temperature.
3. Stop Chlorine flow.
4. Verify return to normal temperature.



HRA Event Tree 2

OMIT TO MONITOR	HEP=0.001	EF=3
FAIL TO DETECT TEMP.RISE	HEP=0.010	EF=3
FAIL TO STOP CHLOR. FLOW	HEP=0.022	EF=3
FAIL TO VERIFY TEMP.	HEP=0.001	EF=3

TOTAL PROBABILITY =0.034

****

Sensitivity Analysis Required? No

END of this Session with THERAX

9.3.5 Fault Tree Analysis

The information and results obtained during the human reliability assessment session above with HERAX, can be input to the system Fault Tree to determine the effects of both human errors and equipment failures on the overall system safety/reliability, as shown in Figure 9.2 and discussed below.

The analysis of the Fault Tree, shown in Figure 9.2, indicates that the expected frequency of a runaway reaction is 0.388 per year. The contribution of operator's errors in not detecting abnormal temperatures or not stopping chlorine flow have the expected probability of 0.034, while the temperature measuring system has the expected probability of failure of 0.31.

The expert system has identified the operator's failure to stop chlorine flow as the most critical task which has the expected probability of 0.022, this failure significantly decreases the operators' overall reliability.

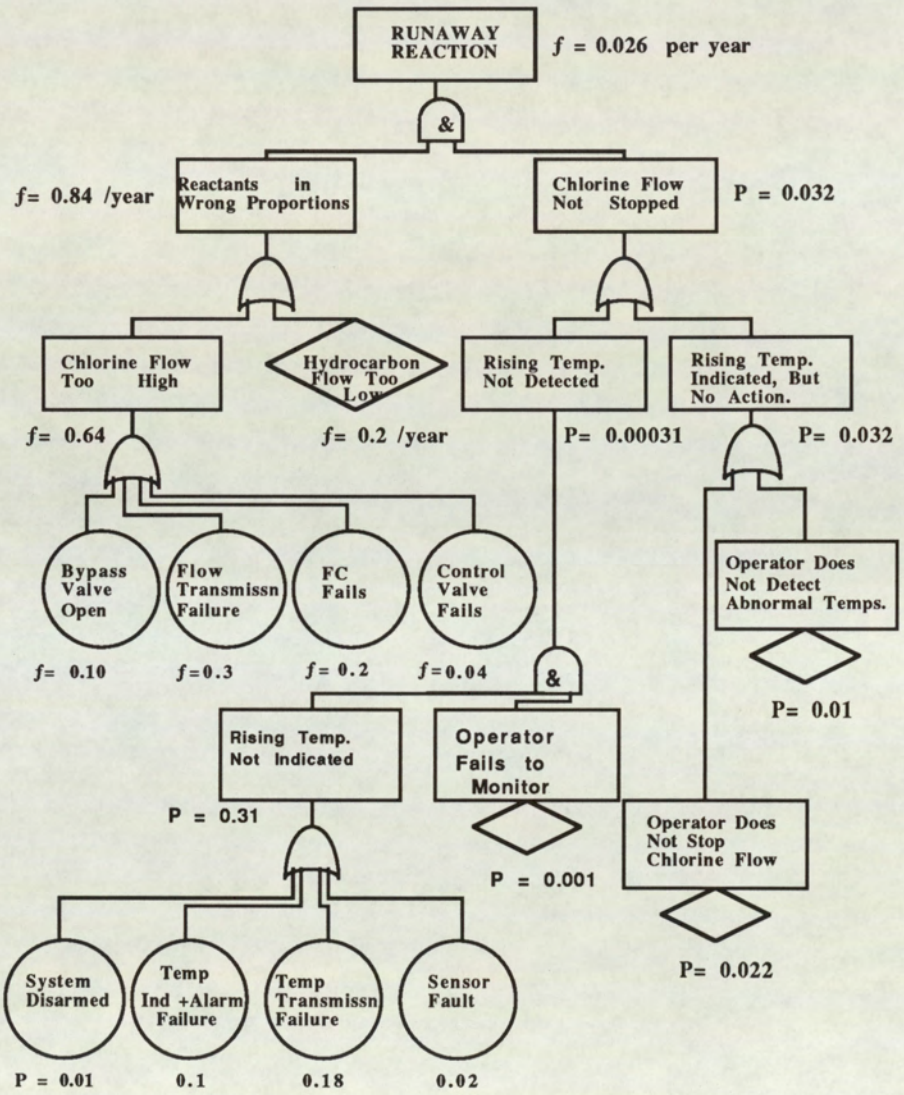


Figure 2 Fault Tree Analysis

9.4 Alternative Design Strategy: *Automatic Control*

As an alternative strategy, the company has decided not to rely entirely upon the process operator to stop the chlorine flow. As a result a trip system was incorporated in the existing process design as shown in Figure 3.

The automatic trip is designed to stop the chlorine flow completely if the temperature reaches 385^{o}K.

It was thought that the new protective system would increase the overall process reliability and reduce the frequency of runaway reactions as a result of human errors.

9.4.1 Objectives of the analysis

The main objective of the analysis for the alternative process design is to determine whether the overall process safety/reliability has improved as a result of adding the automatic protective system to the original process design, i.e., to determine the following:

1. What is the probability that the operator does not detect abnormal rates of temperature rise, or temperature reaching 375^{o}K, knowing that plant policy does not stipulate continuous monitoring of temperature changes?

2. What is the probability that the operator fails to stop the chlorine flow before a runaway reaction occurs ?

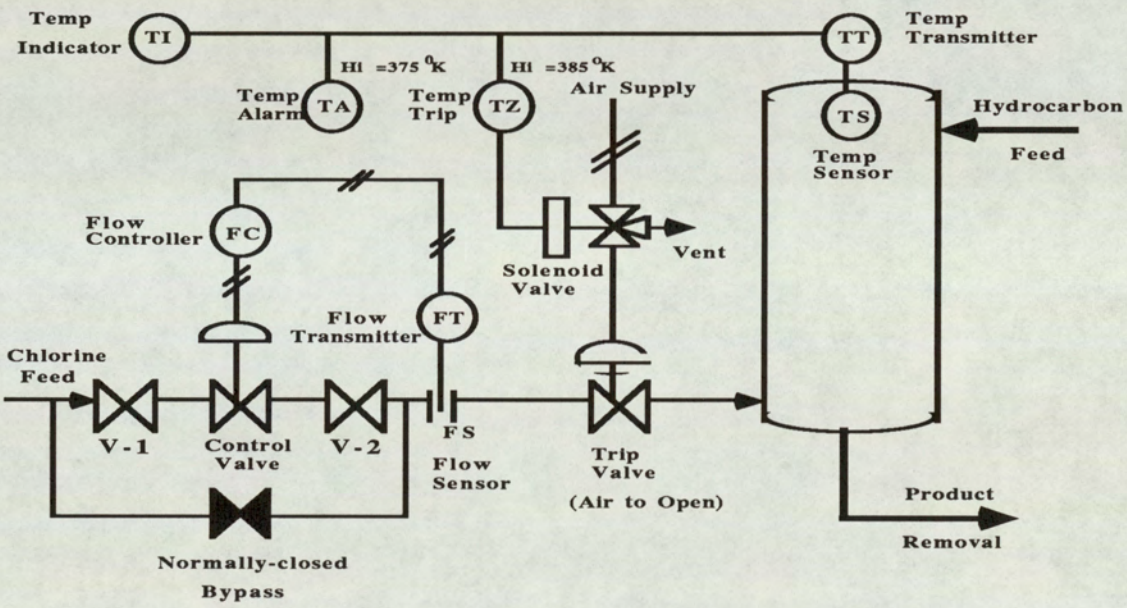


Figure 3 Alternative Process Design

9.4.2 Assumptions

1. The operator is now no longer required to continuously monitor for temperature changes. Therefore, he is expected to devote more of his time attending other needs to keep the plant on-line.

2. Although there are no specific cues for the operator to monitor TI, and most safety-related functions are unannunciated, it is assumed that the operator will carry out hourly scans of the related unannunciated displays in anticipation of abnormal events.

3. In the event of trip system failure to stop the Chlorine flow, it is expected that the operator will anticipate abnormal operating conditions before the trip is activated, but not to recognize that the trip system has failed to operate when required.

4. It is assumed that the only way the operator can prevent runaway reaction is to manually close either V-1 or V-2 that has failed to automatically close.

5. No written procedures exists to assist the operator in the diagnosis of abnormal events.

6. It is estimated that a maximum of 10 minutes is needed for the operator to complete the manual recovery actions to prevent a complete runaway reaction.

7. It is assumed that the operator will have to perform all required actions without any assistance from other plant personnel. Therefore, all HEPs, except for diagnosis, take into account performance of a single operator.

9.4.3 Tasks Identification

The main expected operator tasks in this situation are to:

1. Detect rapidly rising temperature, or high temperature level following the alarm.
2. Recognize/ Diagnose that the trip system has failed.

3. Take appropriate action as early as possible: implement the post-diagnosis Rule-Based actions indicated by the diagnosis "Stop the chlorine flow manually".
4. Return back to the control room after closing either valves V-1 or V-2 and verify the return of temperature to normal.

9.4.4 Expert System Analysis

****

Type of Situation? Abnormal
Task to be analysed? Monitor Temp.
Administrative Control? Yes
Operator Action? Carry out Plant Policy ; (e.g., periodic Testing/Maintenance...)

****

Task being considered:	Monitor Temp.	
FAIL CARRY OUT PLANT POLICY	HEP=0.01	EF=5

****

More Tasks? Yes, Detect Rising Temp.
Type of Actions? Rule-Based
Type of Errors? Commission
M-MI? Displays
Type of Display? UNAN. ; (Table3 and Event Tree1 are displayed)
Operator Action? Scan (hourly)
Type of UNAN? Digital

Task being considered:	Detect Rising Temperature	
FAIL SCAN	HEP = 0.001	EF=10

****

More Actions? No
More Tasks? Yes, Detect Rising Temperature ; (Same task above, but type of
Type of Actions? Diagnosis actions different)
Time Elapsed after Alarm? 10 Min.

Task being considered:	Detect Rising Temperature	
FAIL DIAGNOSE	HEP = 0.1	EF=10

****

Recovery Factors Assessment? Yes

Annunciated Displays

; (Although there are many types of RFs, annunciation of deviant conditions (alarms) are the most attention-getting and therefore, guidance is provided within HERAX to assess their effects. A Table, as well as an Event Tree, are displayed showing Subtasks/PSFs and possible errors associated with the use of Alarms.)

****

Operator Action? Respond

Number of ANNs? 10 ; (10 Alarms)

FAIL NOTICE ONE of 10 ANNs	HEP = 0.05	EF=10

****

More Actions? No

Task being considered:	Detect Rising Temperature	
FAIL SCAN	HEP = 0.001	EF=10
FAIL DIAGNOSE	HEP = 0.05	EF=10
TOTAL FAILURE = 051		

****

More Tasks? Yes, Stop Chlorine Flow

Type of Errors? Omission ; (Omit Initiate Action)

Written Materials Used? Yes

Operator Action? Use Procedural Items

Task being considered:	Stop Chlorine Flow	
OMIT ITEM	HEP = 0.001	EF= 3

****

More Tasks? Yes, Stop Chlorine Flow ; (Same as above, but this time
Type of Errors? Commission errors type is Commission).
M-MI? Manual Valve ; (Table3 above, as well as an
Operator Action? Select (Right valve) Event Tree are displayed)..
Valve Design? Unclearly labeled, part of a group of 2 or more valves similar in ALL the following: size, shape, state, and presence of tags.

Task being considered:	Stop Chlorine Flow	
FAIL SELECT	HEP = 0.01	EF=3

****

More Actions? Yes, Manipulate (Close valve)

Task being considered:	Stop Chlorine Flow
REVERSAL ERROR	HEP = 0.0001

****

More Actions? No

Task being considered:	Stop Chlorine Flow	
FAIL INITIATE ACTION	HEP= 0.001	EF= 3
FAIL SELECT VALVE	HEP= 0.01	EF=3
REVERSAL ERROR	HEP= 0.0001	
TOTAL PROBABILITY = 0.0111		

****

Assessment of PSFs Effect? Yes
Stress Level? Moderately High
Experience Level? Skilled

STOP CHLORINE FLOW	HEP= 0.022	EF=3

; (the HEP is multiplied by a factor of 2)

****

More Tasks? Yes, Verify ; (Temperature back to normal)
Type of Error? Commission

M-MI? UNAN Display

Type of UNAN? Digital Indicator

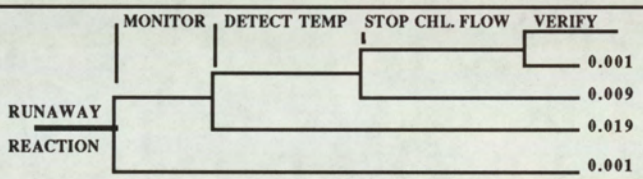
Operator Action? Check-Read

Task being considered:	Verify Temperature	
FAIL CHECK-READ	HEP= 0.001	EF=3

****

More Tasks? No

Tasks being considered		
1. MONITOR		
2. DETECT RISING TEMPERATURE		
3. STOP CHLORINE FLOW		
4. VERIFY TEMPERATURE		



HRA Event Tree 3

OMIT TO MONITOR	HEP=0.01	EF=5
FAIL TO DETECT TEMP.RISE	HEP=0.051	EF=3
FAIL TO STOP CHLOR.FLOW	HEP=0.022	EF=3
FAIL TO VERIFY NORMAL TEMP.	HEP=0.001	EF=3

TOTAL PROBABILITY = 0.084

****

Sensitivity Analysis Required? No

END OF THIS SESSION WITH THERAX

9.4.5 Fault Tree Analysis

Analysis of the Fault Tree shown in Figure 4 indicates that the expected frequency of a runaway reaction is 0.3948 per year. This is marginally greater than in the case of manually operated process. The analysis has shown that the probability of human errors in not detecting abnormal temperatures or not stopping the Chlorine flow has increased to 0.084, and the trip system's **fractional dead time**, based on yearly testing interval, is 0.45 (Kletz, 1984).

The expert system has also identified in this case two critical errors: failure to detect rising temperature and failure to stop chlorine flow, as the most significant operator errors which influence the overall process safety/reliability. The probability of failure to detect rising temperature has increased from 0.01 to 0.51 and the probability of the operator failure to stop chlorine flow has not been reduced.

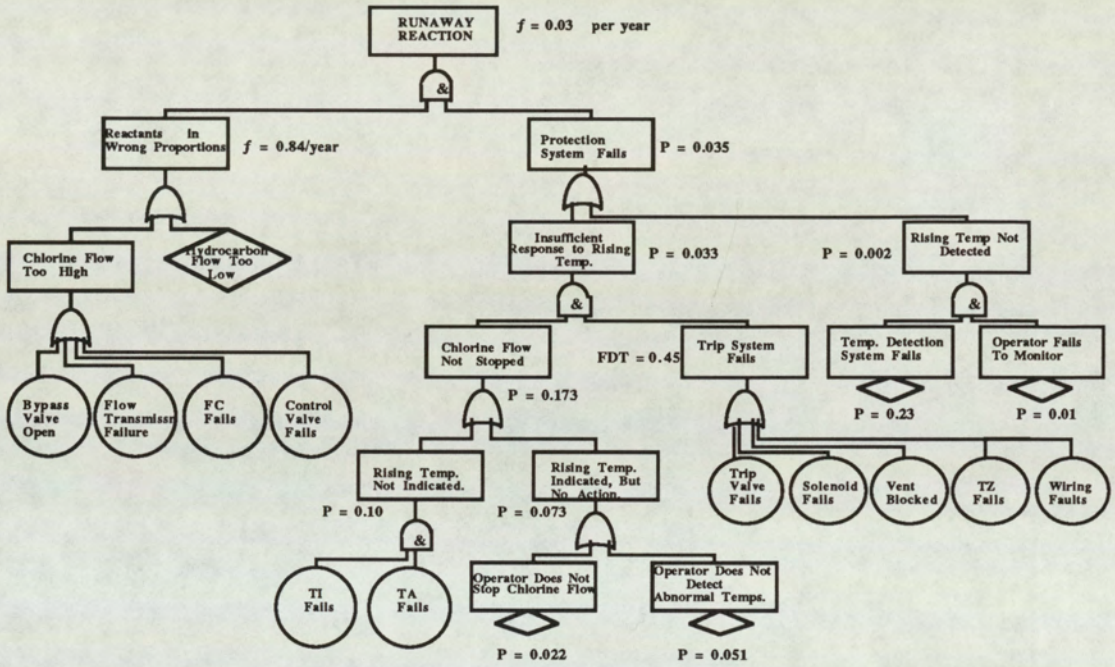


Figure 4 Fault Tree Analysis

9.5 Discussion of Results

The expert system approach described in this Chapter 8 has been applied for the identification and quantification of potential errors in the management of abnormal events. Two different strategies were analysed:

- In the first case, the process is manually controlled and the operator has been instructed to continuously monitor process parameters (the temperature increase) to determine whether the process is operating correctly; therefore, he is more prepared to detect and interpret any parameter deviation from the normal.

- In the second case an alternative strategy was to rely upon an automatic protective system intended to reduce the effects of human errors and was thought to increase the process safety/reliability.

Under the assumptions given, the results show that in the first case the effect on the top event (runaway reaction) of the total operator failure to detect the abnormal temperature and carry out the corrective actions is relatively poor.

The operator performance has been affected by the number of unnecessary alarms and the stress situation, due to the short time allowed for him to implement the written operating procedures and to take the correct course of action.

Analysis of the fault tree shows in the second case that despite adding automatic trip system to the original process design, the operator's performance has greatly deteriorated, particularly in the failure to detect abnormal process conditions. The expected frequency of runaway reaction is also marginally greater than that of the first case.

Reliance on the protective system as well as management system shortcomings in allocating other duties to the operator are chiefly to blame for deterioration in the performance of critical tasks, and the trip system selected was also inherently unreliable.

Automation should be applied with caution and should not be regarded merely as a substitute for the human operator. It often increases human tasks instead of assisting him. An objective and systematic probabilistic risk analysis (PRA) study is, therefore, needed which can take into account the capabilities, limitations and needs of both operators and equipments and help in deciding which functions should be assigned to people, to machines, or to some combination of the two to insure adequate producibility, maintainability, operability, performance, reliability and safety at an acceptable cost.

The expert system for human reliability analysis which has been developed here forms an integral part of such a study and could be a useful tool in aiding the analyst to achieve this objective more easily, more economically and more accurately.

9.6 CONCLUSION

The objective of this chapter was to present a case study to demonstrate the use of the expert system approach HERAX, to assess the relative effects of human reliability following two different (manual versus automatic) strategies on plant safety and reliability.

Although the expert system's methodology and procedures, as described in Chapter 8, have not been fully applied in this simplified case study, it has demonstrated the ability of HERAX to analyse human reliability problems in process plants.

The case study used here is however, an illustration only. A considerable amount of work remains to be done to improve the system in order to be able to analyse more complex process control situations.

The next chapter discusses the general conclusions of this research, the main capabilities and limitations of the expert system approach developed here, and proposes some recommendations for its future improvement.

CHAPTER 10

CONCLUSIONS and RECOMMENDATIONS FOR FUTURE RESEARCH

10.1 Introduction

This thesis has presented an approach, described in Chapter 8, called HERAX, for human reliability modelling and quantification in the process control industries, developed using Artificial Intelligence techniques.

HERAX was developed on an IBM-PC/AT (640 KB RAM) with DOS version 3.0, enhanced graphics, a 20 MB hard disk, and 3 MB extended memory. It uses a mathematic co-processor, and a colour display screen. The system was implemented in GCLisp and took 500 KB.

The two main objectives of this chapter are to present a summary of the conclusions and findings of this thesis and to propose some recommendations for further research and improvement of the system developed here.

10.2 Conclusions

The main conclusions that can be drawn from this research and some of the several fundamental lessons about building an intelligent analyst system that were learned from this project are presented below.

1. One of the most important findings of this thesis is that the analysis of human reliability, as formulated here, is a problem typical of a wide range of health/safety/reliability analysis problems. Knowledge in such domains is mostly

heuristic, residing in the experience of a few human experts, whose skills are in high demand. These characteristics make HRA domain ideal for expert system technology.

2. Production rules have been found quite appropriate for modelling most of the problem-solving strategies carried out by human reliability expert analysts.

3. The development of the HERAX system demonstrates that expert system techniques are applicable and present potential benefits to HRA in process industries.

4. This research project has also shown that it is possible to integrate different HRA techniques in one systematic and easy to use analytical procedure.

5. Although the performance of the HERAX approach is limited by the knowledge acquired from the available literature, the number of rules which were used due to the capacity of the computer, and time constraints, it still demonstrates the usefulness of the approach.

6. While the evaluation exercise, described in Chapter 9, has demonstrated some of the capabilities of the expert system to analyse human reliability in process plants, many questions stay unanswered, for example, its ease of use and acceptability by the intended user population, as well as its accuracy.

7. A major lesson learned from this research work was the need for a domain expert. Although it was possible to build an expert system for HRA, and to integrate different HRA knowledge bases available in the literature using existing expert systems principles (the same concept can be extended to integrate HERAX to existing safety and reliability expert systems), however, had the expert previously identified the characteristics of the domain problem and provided examples of successful cases, development time for this project would have been much shorter.

8. Another critical finding drawn from this research is that development time of a system such as HERAX would have been much shorter if it was designed and developed by a knowledge engineer with AI programming experience and some familiarity with process control industries.

9. The expert system HERAX was developed from scratch, using a computer programming language.However, coding this new system from scratch did not allow concentrating primarily on the knowledge required for high performance. Rather, more time was spend on debugging the procedures that access and manipulate the knowledge.

10. The overall goal of HERAX as a safety/reliability an industrial application of AI techniques has been to make the process of human reliability analysis more easy, cost effective, and accurate. There is a high confidence that a tool such as HERAX can achieve significant safety and reliability improvements and aid in the performance of better analysis. It can do this by making the analysis of HRA more systematic, by aiding engineers in the evaluation of existing or proposed designs.

10.2 Recommendations for Future Research

The expert system approach developed and described in this thesis is only an experimental project for testing purposes, and HERAX is incomplete at this point as a framework for different HRA methodologies and their computational representations to be used by non-experts. The knowledge-based expert system proposed here is not claimed to be suitable for predicting human performance in all circumstances. This system is intended to help engineers and managers in selecting and applying three HRA techniques in process control plants as part of a probabilistic risk assessment. A

considerable amount of work remains to be done and several issues remain unresolved in this work. These issues could constitute potential areas for future research to improve the computer expert system performance and expand its utility.

1. Some of the major factors that should be considered in future building of an analysis expert system are availability, cost, and appropriateness of software and hardware for the scope of the task, as well as availability of the human expert(s) and time necessary to develop and validate the system.

2. There is a need to define more precisely the model and requirements of the potential users of this expert system. It is suggested to extend the system to incorporate inferences made about the user's knowledge, his errors and potential misconceptions to make progress along this line.

3. There is also a requirement to expand and define the knowledge bases accurately and thoroughly by carrying out interviews with experts in the domain to minimize extrapolations. By making the analysis skills of human experts more available to non-experts, the expert system can improve the analysis efficiency.

With respect to this requirement, future program development should emphasize on refining the knowledge base of particular analysis technique and the extension of the system to include other human reliability analysis techniques. At the moment, HERAX is an analysis assistant for the prediction of human reliability using three analytical procedures only, a data bank on human error probabilities (HEPs), which is more applicable to nuclear power plants situations, and a limited list of performance shaping factors (PSFs). HERAX should be upgraded to assist users with a full spectrum of human reliability data and Performance Shaping Factors, associated with specific situations.

4. In the present implementation of HERAX, the Lisp programming environment has made the incremental incorporation of rules easy, but this process is not yet fully systematized or mechanized. Rules are generated through use of a simple text editor. The addition and deletion of rules are also accomplished with this editor. Development of a maintenance system would facilitate the rapid updating of the knowledge base by users as well as experts.

5. Enhancements also should include use of other knowledge representation methods , such as frames and blackboards, and work in the area of the user interface. The production of a more transparent and friendly user interface is a project of importance for the practical application of HERAX

6. There is a need for more test (field) work to improve and validate the performance and accuracy of the system developed in this thesis.

7. Finally, the possibility of integrating HERAX with other safety-related, existing or planned, expert systems should be investigated.

REFERENCES

ACMH (1974), Advisory Committee on Major Hazards.

Adams, J.A.(1982), "Issues in human reliability", Human Factors, No.24, pp.1-10.

Alty, J.L., and Coombs, M.J.(1984), Expert Systems - Concepts and Examples, NCC Publications, Manchester, UK.

Ancelin, P.(1987), "EXPRESS: An Expert System to perform System Safety Studies," Proceedings of an International Topical Conference on Probabilistic Safety Assessment & Risk Management, Swiss Federal Institute of Technology (ETH), Zurich, 30 August-4 September 1987, pp.262-267.

Andow P., and Ferguson G.(1987), "Applications of Knowledge-Based Systems in Chemical Process Safety," World Bank/AIChE/EPA, International Symposium on Preventing Major Chemical Accidents, Washington, DC, pp. 1.129 - 1.143.

Andow, P.(1988), "Safety Integrity Using Expert Systems," Proceedings of the Conference on Expert Systems & Industrial Hazards, IBC, The Cafe Royal , London, 31 October 1988.

Annett, J., and Duncan, K.D.(1967), Task Analysis, Department of Employment, Training Information, Paper 6, HM Stationary Office, London.

ANS (1985), Topical Meeting on Computer Applications for Nuclear Power Plant Operation and Control, Tri-Cities (Pasco), Washington, 8-12 September 1985.

Askren and Regulinski, (1969), "Quantifying Human Performance for Reliability Analysis of Systems." Human Factors, 11(4), 393-396.

Bainbridge, L.(1984), "Process Control," Paper presented at the International Conference on Occupational Ergonomics, 7-9 May 1984, Toronto, Canada.

Bainbridge, L.(1986), "What Should a 'Good' Model of the NPP Operator Contain?." In: ANS/ENS Proceedings of the International Topical Meeting on Advances in

Human Factors in Nuclear Power Systems, Knoxvilee, Tennessee, 21-24 April 1986, pp. 3-11.

Barr, A., and Feigenbaum, E. A.(1981-82), eds., The Handbook of Artificial Intelligence, Vol.I and II, William Kaufmann, Los Altos, California.

Baybutt, P.(1985), "Decision Support Systems and Expert Systems for Risk and Safety Analysis," I.Chem.E Symposium Series, No.97, pp.261-267.

Beare, A.N., and Dorris, R.E.(1983), "A Simulator-based Study of Human Errors in NPP Copntrol Room Tasks, In: Proceedings, Human Factors Society Annual Meeting, pp. 170-174.

Beck, C.F.(1963), "Engineering out the distance factor - A progress report on reactor site criteria," In: Annual Convention of the Federal Bar Association, Philadelphia, Pensilvania, 25 September 1963, (cited in F.R.Farmer (Ed.), Nuclear Reactor Safety, Academic Press, London, 1977.)

Bell, B.J.(1983), "Human Reliability Analysis: A Case Study," In: Low-Probability/High Consequence Risk Analysis, (V.T.Covello and R.A.Waller, eds.), Plenum Press, New York, PP. 297-308.

Bell, B.J., and Swain, A.D.(1981), "Overview of a Procedure for Human Reliability Analysis," In: Proceedings of the International ANS/ENS Topical Meeting on Probabilistic Risk Assessment, Port Chester, NY, 20-24 September 1981, American Nuclear Society, LaGrange Park, IL, pp. 587-596.

Bell, B.J., and Swain, A.D.(1983), A Procedure for Conducting a Human Reliability Analysis for Nuclear Power Plants, Sandia National Laboratories, NUREG/CR-2254, U.S Nuclear Regulatory Commission, Washington, DC, May 1983.

Bellamy, L.(1985), Offshore Drilling Study, (cited in UKAEA, op.cit.).

Bello, G.C., and Colombari, V.(1980), "The Human Factors in Risk Analyses of Process Plants: The Control Room Operator Model, 'TESEO'." In: Reliability Engineering, No. 1, pp 3-14.

Beltracchi, L.(1988), "An Expert Display System and Nuclear Power Plant Control rooms," IEEE Transactions on Nuclear Science, Vol. 35, No. 2, April 1988.

Berliner, D.C., Angell, D, and Shearer, J.W.(1964), "Behaviours, Measures, and Instruments for Performance Evaluation in Simulated Environments," pp. 275-296, in H.R. Leuba (ed) Proceedings of the Symposium on the Quantification of Human Performance, Electronics Industries Assn. and University of New Mexico, Albuquerque, NM, August 1964.

Blanchard, R.E, Mitchell, M.B., and Smith, R.L.(1966), Likelihood of Accomplishement Scale for a sample of Man-Machine Activities, Dunlap and Associates, Inc., Santa Monica, CA, June 1966.

Blix, H.(1988), "Opening Remarks", Proceedings of an International Conference on Man-Machine Interface in the Nuclear Industry, IAEA, Vienna.

Brown, J.S, Burton, R.R., and deKleer, J.(1982), "Knowledge engineering and pedagogical techniques in SOPHIE I, II and III". In: Intelligent Tutoring Systems (D. Sleeman and J.S. Brown, eds.), pp. 227-282, Academic Press, New York.

Brune, R.L., Weistein, M., and Fitzwater, M.E.(1983), Peer Review Study of the Draft Handbook for Human Reliability Analysis With Emphasis on Nuclear Power Plants Applications, NUREG/CR-1278, Human Performance Technologies, Inc. SAND82-7056, Sandia National Laboratories, Albuquerque, NM, January 1983.

Buchanan, B.G., and Feigenbaum, E.A., "DENDRAL and Meta-DENDRAL: Their applications dimensions," Artificial Intelligence, 11, 1978, pp.5-24.

Buchanan, B.G., and Shortliffe, H.E.(1984), Rule-Base Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project, Reading, MA: Addison-Wesley.

Carlesso, S., Barbas, T., Capobianchi, A., Koletsos, A., Mancini, G.(1987), "Application of Artificial Intelligence to Reliability Data Banks," Proceedings of an International Topical Conference on Probabilistic Safety Assessment & Risk

Beltracchi, L.(1988), "An Expert Display System and Nuclear Power Plant Control rooms," IEEE Transactions on Nuclear Science, Vol. 35, No. 2, April 1988.

Berliner, D.C., Angell, D, and Shearer, J.W.(1964), "Behaviours, Measures, and Instruments for Performance Evaluation in Simulated Environments," pp. 275-296, in H.R. Leuba (ed) Proceedings of the Symposium on the Quantification of Human Performance, Electronics Industries Assn. and University of New Mexico, Albuquerque, NM, August 1964.

Blanchard, R.E, Mitchell, M.B., and Smith, R.L.(1966), Likelihood of Accomplishement Scale for a sample of Man-Machine Activities, Dunlap and Associates, Inc., Santa Monica, CA, June 1966.

Blix, H.(1988), "Opening Remarks", Proceedings of an International Conference on Man-Machine Interface in the Nuclear Industry, IAEA, Vienna.

Brown, J.S, Burton, R.R., and deKleer, J.(1982), "Knowledge engineering and pedagogical techniques in SOPHIE I, II and III". In: Intelligent Tutoring Systems (D. Sleeman and J.S. Brown, eds.), pp. 227-282, Academic Press, New York.

Brune, R.L., Weistein, M., and Fitzwater, M.E.(1983), Peer Review Study of the Draft Handbook for Human Reliability Analysis With Emphasis on Nuclear Power Plants Applications, NUREG/CR-1278, Human Performance Technologies, Inc. SAND82-7056, Sandia National Laboratories, Albuquerque, NM, January 1983.

Buchanan, B.G., and Feigenbaum, E.A., "DENDRAL and Meta-DENDRAL: Their applications dimensions," Artificial Intelligence, 11, 1978, pp.5-24.

Buchanan, B.G., and Shortliffe, H.E.(1984), Rule-Base Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project, Reading, MA: Addison-Wesley.

Carlesso, S., Barbas, T., Capobianchi, A., Koletsos, A., Mancini, G.(1987), "Application of Artificial Intelligence to Reliability Data Banks," Proceedings of an International Topical Conference on Probabilistic Safety Assessment & Risk

Management, Swiss Federal Institute of Technology (ETH), Zurich, 30 August-4 September 1987, pp.287-292.

Carlson, D.D., Gallup, D.R., Kolaczkowski, A.M., Kolb, G.J., Stack, D.W., Horton, W.H., and Lobner, P.R.(1983), Interim Reliability Evaluation Program Procedures Guide, Sandia National Laboratories, NUREG/CR-2728, U.S. Nuclear Regulatory Commission, Washington, DC, January 1983.

Carson, P.A, and Mumford, C.J.(1988), The Safe Handling of Chemicals in Industry, Vol.2, Longman & Technical, Essex, England.

Chambers, A.N.(1969), Development of a Taxonomy of Human Performance: A Heuristic Model for the Development of Classification Systems, AIR-7263/69-TR-4A, American Institute for Research, Washington, DC, August 1969, (cited in A.D.Swain and H.E.Guttmann, op.cit).

Chapanis, A.(1965), Man-Machine Engineering, Belmont, CA: Wadsworth Publishing Co.

Chatfield, C.(1978), Statistics for Technology, 2nd ed, Publ. Chapman & Hall, London.

Chemical Industries Association (1977), A Guide to Hazard and Operability Studies, London: Alembic House.

CIMAH (1984), Control of Industrial Major Accidents Hazard Regulations.

Cmnd. 6618 Royal Commission on Environmental Pollution (1986), Sixth Report, Nuclear Power and the Environment, HMSO, London, 1976, pp. 112-115 and p. 202, (cited in J.C. Consultancy Ltd., Risk Assessment for Hazardous Installation, London, 1986.)

Colley, R.(1985), "Review of Artificial Intelligence for Nuclear Applications," Transactions of the American Nuclear Society, Winter Meeting, S.F., California, 10-14 November 1985, pp.291-299.

Comer, M.K, Kozinsky, E.J., Eckel, J.S., and Miller, D.P.(1983), Human Reliability Data bank for Nuclear Power Plant Operations, Volume 2: A Data

Bank Concept and System Description, General Physics Corporation and Sandia National Laboratories, NUREG/CR-2744, U.S. Nuclear Regulatory Commission, Washington, DC, February 1983.

Comer, M.K., Seaver, D.A., Stillwell, W.G., and Gaddy, C.D.(1984), Generating Human Reliability Estimates Using Expert Judgement, NUREG/CR-3688, Sandia National Laboratory, Albuquerque, New Mexico 87185, USA.

Conference Record (1982), the 1981 IEEE Standards Workshop on Human Factors and Nuclear safety, IEEE.

Cox, R.A., and Slater, D.H.(1984), "State-of-the-Art, of Risk Assessment of Chemical Plants in Europe." In: Low-Probability High-Consequences Risk Analysis, (R. A. Waller , and V. T. Covello, eds.), Plenum Press, New York, pp. 257-283.

Dalkey, N.C.(1969), The Delphi Method: An Experimental Study of Group Opinion, Memorandum, RM-5888-PR, June 1969, RAND Corporation, Santa Monica, California, (cited in UKAEA, op.cit).

Dhillon, B.S.(1987), Human Reliability With Human Factors, Pergamon Press, U.K.

Dixon B.W., and Hinton M.F.(1985), "Reviewing the Development of an Artificial Intelligence Based Risk program," Transactions of the American Nuclear Society, Winter Meeting, S.F., California, 10-14 November 1985.

Dougherty, E.M.(1983), "Survey of How PRAs Model Human Error," In: Proceedings of the ANS/ENS International Meeting on Thermal Nuclear Safety, 29 August-2 September 1983, Chicago, IL, American Nuclear Society, NUREG/CP-0027, U.S. Nuclear Regulatory Commission, Washington, DC, January 1983, pp. 565-574.

Dougherty, E.M., and Fragola, J.R.(1988), Human Reliability Analysis: A Systems Engineering Approach With Nuclear Power Plant Applications, John Wiley & Sons, U.S.A.

Drury, C.G, Paramore, B., Van Cott, H.P., Grey, S.M., and Corlett, E.N.(1987), "Task Analysis." In: Handbook of Human Factors, G.Salvendy, (ed), John-Wiley & Sons, New York, pp. 370-401.

Duda, R.O., Gaschning, J.G., Hart, P.E.(1979), "Model design in the Prospector consultant system for mineral exploration". In: Expert Systems in the Micro-Electronic Age, (D. Michie, ed.), Edinburgh University Press, Edinburgh, pp.153-167.

Duncan, K.D.(1974), "An Analytical Technique for Industrial Training." In: W.T.Singleton and P.Spurgeon, (eds), Measurement of Human Resources, Taylor and Francis, London.

EDF-DER (1984), Un Langage pour les systems Experts: ALOUETTE, EDF-DER, Note HI/4773, (Cited in P. Ancelin, op.cit.).

EDF-DER (1985), "Notice d'utilisation et analyse des logiciels L.R.C." EDF-DER, Note HI/5097-02, (Cited in Ancelin, op.cit.).

Edwards, E.M.(1984), "Recent Developments in Quantifying Human Performance," Transactions of the ANS/ENS, pp. 207-208.

Edwards, W.(1977), "How to Use Multiattribute Utility Measurement for Social Decision Making," In: IEEE Transactions on Systems, Man, and Cybernetics, SMC-7, pp.326-340.

EEC's "Seveso" Directive of 1982 on Major-Accidents Hazards (82/501/EEC), Official Journal of the European Communities, 5:8:82, l230/1 Vol. 25.

Embrey, D.E (1983), The Use of Performance Shaping Factors and Quantified Expert Judgement in the Evaluation of Human Reliability: An Initial Appraisal, NUREG/CR-2986, Brookhaven National Laboratory.

Embrey, D.E, Humphreys, P., Rosa, E.A., Kirwan, B., and Rea, K.(1984), SLIM-MAUD: An Approach to Assessing Human Error Probabilities Using Structured Expert Judgement," Vol. I and II, NUREG/CR-3518, Brookhaven National Laboratory, Upton, New York 11973, USA.

Embrey, D.E.(1976), Human Reliability in Complex Systems: An Overview Report R10, NCSR, UKAEA, Wigshaw, Lane, Warrington, WA3 4NW, England.

Embrey, D.E.(1981), "A New Approach to the Evaluation and Qualification of Human Reliability in Systems Assessment," In: Proceedings fo the Third National reliability Conference - Reliabity 81, Birmingham, England, pp. 5B/1/1-5B/1/12.

Embrey, D.E.(1981), "Approaches to the Evaluation and Reduction of Human Error in the Process Industries." In: IChemE Symposium Series, No. 66.

Embrey, D.E.(1987), "Human Reliability," Proceedings of a Conference on Human Reliability in Nuclear Power Plants, IBC, London, pp. 1-35.

Embrey, D.E., Humphreys, P., Rosa, E.A., Kirwan, B., and Rea, K.(1984), SLIM-MAUD: An Approahc to Assessing Human error Probabilities Using Structured Expert Judgement. Vol. I and II. NUREG/CR-3518. Brookhaven National Laboratory, Upton, New York 11973, USA.

Evans, R.A.(1976), "Reliability Optimisation," pp. 117-131, in E.J. Henley and J.W. Lynn (eds), Generic Techniques in Systems Reliability Assessment, Leyden, The Netherlands: Noordhoff International Publishing, (cited in A.D Swain and H.E.Guttmann, op. cit.).

Farmer, F.R.(1964), "The growth of reactor safety criteria in the United Kingdom," Anglo-Spanish Nuclear Power Symposium, Madrid, November 1964, (cited in F.R.Farmer (Ed.), Nuclear Reactor Safety, Academic Press, London, 1977.)

Farmer, F.R.(1967), Siting Criteria, a New Approach, IAEA SM-89/34, Vienna.

Farr, D.E.(1984), "Human Factors in the Nuclear Industry - A Critical Review," In: Proceedings of the 1984 International Conference on Occupational Ergonomics, pp. 11-16.

Feigenbaum, E.A.(1977), "The Art of of Artificial Intelligence: Themes and Case Studies in Knowledge Engineering," Proceedings IJCAI-77, pp.1014-1029.

Fleishman, E.A., Kincade, R.G., and Chambers, A.N.(1968, 1970), Development of a Taxonomy of Human Performance, American Institutes for Research, Washington, DC, (cited in Swain and Guttmann op.cit.).

Flight International, No.22, January 1975.

Fu, K.S, Gonzales, R.C., Lee, G.C.G.(1987), Robotics: Control, Sensing, Vision and Intelligence, McGraw-Hill, New York.

Gallacher, J.(1989), "Practical introduction to expert systems," In: Microprocessors and Microsystems, Vol 13, No. 1, January/February 1989, pp. 47-53.

Garriba, S.F., Guagnini, E., Mussio, P.(1987), "Knowledge Paradigms and Error Management in an Expert System for Fault-Tree Construction," Proceedings of an International Topical Conference on Probabilistic Safety Assessment & Risk Management, Swiss Federal Institute of Technology (ETH), Zurich, 30 August-4 September 1987, pp.268-275.

Gifford, C.(1989), "Hinkley Point C inquiry report, Risk tolerability," Atom, No. 392, June 1989, pp. 27-28.

Green, A.E., and Bourne, A.J. (1974), Reliability Technology, London: John Wiley.

Hagen, E.W., and Mays, G.T.(1981), "Human Factors Engineering in the U.S. Nuclear Arena," Nuclear Safety, Vol. 22, No. 3, May-June 1981, pp. 337-345.

Hale, A.R., and Glenden, A.I.(1987), (eds), Individual Behaviour in the Control of Danger, Elsevier Science Publishers, Amsterdam, The Netherland.

Hall, R.E.,, Fragola, J., and Wreathall, J.(1982), Post Event Human Decision Errors: Operator Action Tree/ Time Reliability Correlation, Brookhaven National Laboratory, NUREG/CR-3010, U.S. Nuclear Regulatory Commission, Washington, DC, November 1982.

Hannaman, G.W., Spurgin, A.J., and Fragola, J.R.(1984b), Systematic Human Action Relaibility Procedure (SHARP), Interim Report, NP-3583, Elewctric Power Research Institute, June 1984.

Hannaman, G.W., Spurgin, A.J., and Lukic, Y.D.(1984a), Human Cognitive Reliability Model for PRA Analysis, NUS-4531, NUS Corporation, December 1984.

Harmon, P., and King, C.D.(1985), Artificial Intelligence in Business, John Wiley & Sons, New York.

Harmon, P.and Maus, R., William, (1988), Expert Systems : tools and applications, Wiley & Sons, New York.

HASAWA (1974), Health and Safety at Work etc. Act .

Hayes-Roth, F, Waterman, D.A, and Lenat, D.(1983), Building Expert Systems, Reading, Massachussetts: Addison-Wesley.

Hayes-Roth, F.(1984), "The Knowledge-Based Expert System: A Tutorial," IEEE Computer, September 1984, pp.11-28.

Hayes-Roth, F.(1985), "Rule-Based Systems". In: Communication of the ACM, Special Session on Architecture for Knowledge-based Systems, Vol. 28, No. 9, September 1985.

Heslinga, G.(1983), "Human Reliability Analysis using Event Trees." In: Kema Scientific and Technical Reports 1(3): 19-44.

Hinton, C.(1957), Axel Axson Johnson Lecture, Stockholm, (cited in J.R.Thomson, op. cit.).

HMSO (1972), Report on Safety and Health at Work 1972 by a committee chaired by Lord Robens, HMSO, London.

Holden, P.L.(1985), "Developments in Risk Assessment and its Applications in process Plant Safety Studies," EFCE Publication, The Assessment and Control of Major Hazards, Series No. 42, Manchester 22-24 April 1985, pp. 263-271.

Hollnagel, E.(1987), "Applications of Knowledge-Based Systems in NPP Control Rooms," Proceedings of Conference on Human Reliability in Nuclear Power, IBC, Regent Crest Hotel, London, 22nd-23rd October, 1987, pp.139-155.

HSE (1976), The Control of Major Hazards, Advisory Committee, First Report, HMSO, London, 1976.

HSE (1979a), Safety assessment principles for nuclear power reactors. HSE - Report HA5, HMSO.

HSE (1979b), The Control of Major Hazards, Advisory Committee, Second Report, HMSO, London, 1979.

HSE (1982), Ergonomics/human factors in the design and safe operation of pressurised water reactors. Report NII - ONSWG (82), (cited in D.Whitfield 1987a, op. cit.).

HSE (1983a), Safety assessment for nuclear chemical plants. Health and Safety Executive.

HSE (1984), The Control of Major Hazards, Advisory Committee, Third Report, HMSO, London, 1984.

HSE (1986), HM Nuclear Installations Inspectorate: Safety Audit of BNFL Sellafield, London: HMSO, (cited in D. Whitefield (1987b) op. cit.).

HSE (1987), The tolerability of risk from nuclear power stations, HMSO, London.

HSE (1989a), Quantified risk assessment: Its input to decision making HMSO, London.

HSE (1989b), Risk criteria for land-use planning in the vicinity of major industrial hazards, HMSO, London.

HSE (1989c), Human factors in industrial safety, HMSO, London.

Hu, D.(1988), Programmer's Reference Guide to Expert Systems, Howard W. Sams & Company, Indiana, USA.

Human Factors (1964), Volume 6, Summary of Papers Presented at a Symposium held by the Electronics Industries Association - Human Factors Subcommitee, Albuquerque, NM, December 1964.

Hushon, J.M.(1986), "Response to chemical emergencies," Environmental Science and Technology, Vol.20, No.2, pp.118-121.

IAEA (1988), Proceedings of an International Conference on Man-Machine Interface in the Nuclear Industry, IAEA, Vienna.

IBC (1988), Proceedings of Conference on Expert Systems & Industrial Hazards, The Cafe Royal , London, 31 October 1988.

IEEE (1981), Report of a Survey for Models and databases Relating to Human Performance in Nuclear Power Generating Stations (working draft, revision 1). Task Group on Human Performance Evaluation of IEEE Human Factors Working Group, SC5,5, Washington, DC, July 1981.

Irwin, I.A., Levitz, J.J., and Ford, A.A. (1964), "Human Reliability in the Performance of Maintenance," pp. 143-148 In: H.R.Lueba (ed), Proceedings of the Symposium on Quantification of Human Performance, Electronics Industries Assn. and Univ. of New Mexico, Albuquerque, NM, August 1964.

Jackson, P.(1986), "Review of Knowledg-representation tools and techniques," In: IEEE Proceedings, Vol. 134, No. 4, July 1986, pp. 224-230.

Kahneman, D., and Tversky, A.(1979), Intuitive Prediction: Biases and Corrective Procedures, TIMS Studies in Management Sciences, 12, Wheelwright and Maridakis, pp. 313-327, (cited in UKAEA, op.cit.).

Kemeny, (1979), Three Mile Island: A Report to the Commissioners and to the Public, Washington, DC.

Kibblewhite, J.(1988), Microcomputer Applications in Safety Management, Reading, U.K., December 1988.

Kirwan, B.(1982), An Evaluation of and Comparison of Three Subjective Human Reliability Quantification Techniques, M.Sc. Dissertation, University of Birmingham, Dept. of Eng. Prod., September (unpublished).

Kletz, T.(1980), "Plant Instruments: Which Ones Don't Work and Why'" Chemical Engineering Progress, July 1980, pp. 68-71.

Kletz, T.(1988), "The Past and Future of Loss Prevention," The Chemical Engineer Centenary Supplement, pp. 25-28.

Klueh, R.(1986), "Future nuclear reactors - safety first?," New Scientist, 3 April 1986, pp. 41-45.

Kolb, G.J., Berry, D.L., Easterling, R.G., Hickman, J.W., Kolaczkowski, A.M., Swain, A.D., Von Riesemann, W.A., Woodfin, R.L., Reed, J.W., McCann, M.W., and Junsman, D.M.(1982), Review and Evaluation of the Indian Point Probabilistic Safety Study, Sandia National Laboratories, NUREG/CR-2934, U.S. Nuclear Regulatory Commission, Washington, DC, December 1982.

Kroger, W., Camarinopoulos, and Caisley, J.(1987), "Critical Review of Analytical Techniques for Risk Studies in Nuclear Power Installations," Proceedings of an International Topical Conference on Probabilistic Safety Assessment & Risk Management, Swiss Federal Institute of Technology (ETH), Zurich, 30 August- 4 September 1987, pp. 24-30.

Lederman, L., and Gubler, R.(1987), "International Experience in PSA: the IAEA Perspective," Proceedings of an International Topical Conference on Probabilistic Safety Assessment & Risk Management, Swiss Federal Institute of Technology (ETH), Zurich, 30 August- 4 September 1987, pp. 16-23.

Lee, K.(1988), "A Literature Survey of the Human Reliability Component in an Man-Machine System," IEEE Transactions on Reliability, Vol.37, No. 1, April 1988.

Lees, F.P.(1980), Loss Prevention in the Process Industries, Vol. 1, London: Butterworths.

Leplat, J.(1985), Erreur Humaine, Fiabilité Humaine dans le Travail, Armand Colin, Paris, 1985.

Leplat, J., and Rasmussen, J.(1984), "Analysis of Human Errors in Industrial Incidents and Accidents for Improvements of Work Safety." In: Accident Analysis and Prevention, Vol. 16, No.8, pp. 77-87

Lindsay, R., Buchanan, B.G., Feigenbaum, E.A., and Lederberg, J.(1980), Applications of Artificial Intelligence for Organic Chemistry: The DENDRAL Project, McGraw-Hill, New York.

Marshall, V.C.(1987), Major Chemical Hazards, Ellis Horwood Ltd., Chichester.

Mc Dermot, F.(1982), "R1: A Rule-based configurer of computer system," Artificial Intelligence, Vol. 19, pp. 39-88.

McCarthy, J.(1978), History of LISP," in: ACM SIGPLAN Notices, Vol.13, No. 8, Also in: History of Programming Languages, Wexelblat, R.L.(Ed.), Academic Press, New York.

McCormick, E.J., and Senders, M.S.(1982), Human Factors in Engineering and Design (5th ed), New York: McGraw-hill.

Meddis, R.(1973), Elementary Analysis of Variance for the Behavioural Sciences, London: McGraw-Hill.

Meister, D.(1966), "Applications of Human Reliability to the Production Process". In W.B.Askren (ed) (1967), Report No. AMRL-TR-67-88, Symposium on Reliability of Human Performance in Work.

Meister, D.(1967), "Tables for Predicting the Operational Performance of Personnel," Appendix A, In: J.J.Hornyak, Effectiveness of Display Subsystem Measurement and Prediction Techniques, September 1967, (cited in Swain and Guttmann op. cit.).

Meister, D.(1971), Comparative Analysis of Human reliability Models, Bunker Ramo Electronics Systems Division, Westlake Village, CA, Report L0074-107, November 1971.

Meister, D.(1982), "Where and What are the Data in Human Factors?" In: Proceedings of the Human Factors Society, 26th Annual Meeting.

Meister, D.(1984), "Human Reliability." In: Human Factors Review: 1984, Frederick A. Muckler, Alan S.Neal, and Lynn Strother (eds), The Human Factors Society, Santa Monica, California, pp.13-53.

Meister, D.(1985), Behavioural Analysis and Measurement Methods.

Meister, D., and Rabideau, G.F.(1965), Human Factors Evaluation in System Development, John Wiley & Sons, New York.

MIL-HDBK-217/Notice1 (1986), Reliability Prediction of Electronic Components, (cited in D.Zemva, et al., op.cit.).

Miller, D.P., and Swain, A.D.(1987), "Human Error and Human Reliability," In: G.Salvendy (ed), Handbook on Human Factors, pp. 219-250.

Miller, D.S.(1988), "Pressure Relief valve Selection, a Collaborative Expert System Approach," Proceedings of the Conference on Expert Systems & Industrial Hazards, IBC, The Cafe Royal , London, 31 October 1988.

Miller, R.A., et al.(1982), "INTERNIST-1, an experimental computer-based diagnostic consultant for general internal medecine," New England Journal of Medecine, 307, 8 , August 1982.

Milne, R.W.(1988), "Rapid Fault Diagnosis," Proceedings of the Conference on Expert Systems & Industrial Hazards, IBC, The Cafe Royal , London, 31 October 1988.

Munger, S.J., SmithR.W., and Payne, D.(1962), An index of electronic equipment operability:Data Store, Report AIR-C43-1/62-RP (1), American Institute for Research, Pittsburgh, PA, January 1962, (AD 607 161).

Murphy, A.H., and Winkler, R.L.(1974), Credible Interval Temperature Forecasting: Some Experimental Results, Monthly Weather Review, No.102, pp.784-794, (cited in UKAEA, op.cit.).

NIHHS (1982), Notification of Installations Handling Hazardous Substances Regulations.

Norman, D.A.(1981), Categorization of action slips, Psychological Review, No.88, pp.1-15.

NUREG-75/014, WASH-1400, Reactor Safety Study - An Assessment of Accident Risks in U.S. Commercial Nuclear Power Plants, U.S. Nuclear Regulatory Commission, Washington, DC, October 1975.

NUREG-1050 (1984), Probabilistic Risk Assessment (PRA) -- Status Report and Guidance for Regulatory Application. Draft Report for Comment, U.S. Nuclear Regulatory Commission, Washington, DC, February 1984.

NUREG/CR-2300 (1983), PRA Procedure Guide, Vols. 1 and 2, American Nuclear Society Institute for Electrical and Electronic Engineers, U.S. Nuclear Regulatory Commission, Washington, DC, January 1983.

NUREG/CR-2300 (1983), PRA Procedure Guide, Vols.1 and 2, American Nuclear Society IEEE, U.S. Nuclear Regulatory Commission, Washington, DC, January 1983.

Okrent, D.(1947), Nuclear Reactor Safety on the History of the Regulatory Process, (cited in P. Tanguy, op. cit.).

Pantony, M.F., Scilly, N.F., and Barton, J.A.(1989), "Safety of Exothermic Reactions: A UK Strategy," In: Plant/Operations Progress, Vol. 8, No. 2, pp.113-117.

Pearson and Brazendale, (1988), " ," World Bank/AIChE, International Symposium on Preventing Major Chemical and Related Process Accidents, London, 10-12 May 1988, pp.195-208.

Perrow, C.(1984), Normal Accidents, NY: Basic Books.

Pew, R.W., and Baron, S.(1983), "Perspectives on Human Performance Modelling," Automatica, Vol.19, No.663.

Pew, R.W., Feehrer, C.E., and Baron, S.(1977), Critical Review and Analysis of Performance Models Applicable to Man-Machine Systems Evaluation, AFOSR-TR-77-0520, Air Force Office of Scientific Research, Bolling AFB, Washington, DC, March 1977.

Phillips, L.D, Humphreys, P., and Embrey, D.E.(1985), "Appendix D: A Socio-Technical Approach to Assessing Human Reliability (STAHR)," In: Selby, D., Pressurized Thermal Shock Evaluation of the Calvert Cliffs Unit 1 Nuclear Power Plant, Research Report on DOE Contract 105840R21400, Oak Ridge National Laboratory, Oack Ridge, TN.

Piso, E.(1981), "Task analysis for process-control tasks: The method of Annett, et al., applied," Journal of Occupational Psychology, 54, pp. 247-254.

Popchev, I, and Zlatareva, N.(1985), "An Expert System in Reliability-Structure and Knowledge Representation. In: Artificial Intelligence: Methodology, Systems, Applications (W. Bible and B. Petkoff, eds.), Elsevier Science Publishers, North-Holland, pp.199-205.

Prerau, D.S.(1985), "Selection of an appropriate domain for an expert system," The AI Magazine, Vol. 6, No. 2, pp. 26-30.

Proceedings (1975), The 1975 Annual Reliability and Maintainability Symposium, January, 1975.

PWR, HSE (1979), A Report by the HSE to the Secretary of State for Energy on a review of the generic safety issues of pressurised water reactors. HMSO, London.

Raafat, H.M.N., and Abdouni, A.H.(1987), "Development of an Expert System for Human Reliability Analysis," Journal of Occupational Accidents. No. 9, pp.137-157.

Rasmussen, J. (1988), " ," World Bank/AIChE, International Symposium on Preventing Major Chemical and Related Process Accidents, London, 10-12 May 1988, pp.535-551.

Rasmussen, J.(1980), "What can be learned from error reports? In: K.D.Duncan, M.M.Grunberg, and D.Wallis, (eds), Changes in Working Life, Wiley: New York.Riso National Laboratory in Denmark.

Rasmussen, J.(1982), "Human Errors. A Taxonomy for Describing Human Malfunction in Industrial Installations." In: Journal of Occupational Accidents, 4, pp. 311-333.

Rasmussen, J.(1986), Information Processing and Human-Machine Interaction: An Approach to Cognitive Engineering, North-Holland, Elsevier Science Publishers, New York.

Rasmussen, J., Carnino, A., Griffon, M., Mancini, G., and Gagnolet, P.(1981), Classification System for Reporting Events Evolving Human Malfunctions, RISO-M-2240, RISO National Laboratory, Denmark, March 1981.

Rasmussen, J., Duncan, K., and Leplat, J.(1987), (eds),.New Technology and Human Error, John Wiley & Sons.

Reason, J and Mycielska, K.(1982), Absent Minded? The Psychology of Mental Lapses and Errors, Prentice-Hall: Englewood Cliffs, New Jersey.

Reason, J.(1987a), "The Psychology of Mistakes: A Brief Review of Planning Failures." In: New Technology and Human Error, J.Rasmussen, K.Duncan, and J.Leplat, (eds), John Wiley & Sons, pp.45-52.

Reason, J.(1987b), "Generic Error-Modelling System (GEMS): A Cognitive Framework for Locating Common Human Error Forms." In: New Technology and Human Error, J.Rasmussen, K.Duncan, and J.Leplat, (eds), John Wiley & Sons, pp.63-83.

Reason, J., and Lucas, D.(1984), "Using cognitive diaries to investigate naturally occuring memory blocks." In: J.Harrisa, and P.Morris (eds), Everyday Memory, Actions and Absent-Mindedness. Academic Press, London.

Reason, J.T.(1987), "The Human Contribution to Nuclear Power Plant Emergencies," Proceeding of a Conference on Human Reliability in Nuclear Power, IBC, London, pp. 110-118.

Reeves, A.B., Linkens, D.A., Wells, G.L.(1988), "The Use of PROLOG for the Development of Expert Checklists for the Identification of Failure Events on Chemical Plant," Proceedings of Conference on Expert Systems & Industrial Hazards, IBC, The Cafe Royal , London, 31 October 1988.

Regulinski, (1973), "On Modelling Human Performance Reliability." IEEE Transactions on Reliability, R-22 (3), 114-115.

Regulinski, T.L.(1973), (Ed), "Special issue on Human Reliability", IEEE Transactions on Reliability, No.22, August 1973.

Regulinski, T.L., and Askren, W.B.(1969), "Mathematical Modelling of Human Performance Reliability." Proceedings 1969, Annual Symposium on Reliability, IEEE, Cat. No. 69-08-R, PP. 5-11.

Regulinski, T.L., and Askren, W.B.(1972), "Stochastic Modelling of Human Performance effectiveness functions." Proceedings 1972, Annual Symposium on Reliability and Maintainability, pp.407-416.

Rigby, L.V.(1970), "The Nature of Human Error," pp. 457-466 In: Annual Technical Conference Transactions of the ASQC, American Society for Quality Control, Milwaukee, WI, May 1970 (also pp. 712-718 In: Chem. Tech, December 1971).

Rook, L.W.(1962), Reduction of Human Error in Industrial Production, Report SCTM, 93-62 (14), Sandia National Laboratories, Albuquerque, NM, June 1962.

Rook, L.W.(1965), Motivation and Human Error, Report SCIM, 65-135, Sandia National Laboratories, Albuquerque, NM, September 1965.

Rosa, E.A., Humphreys, P.C., Spettell, C.M., and Embrey, D.E.(1985), Application of SLIM-MAUD: A Test of an Interactive Computer-Based Method for Organising Expert Assessment of Human Performance and Reliability, NUREG/CR-4016, Brookhaven National Laboratory, Upton, New York, for Office of Nuclear Regulatory Commission, Washington, CD 20555, September 1985.

Schmall, T.(1980), (ed), Conference Record for 1979 IEEE Standards Workshop on Human Factors and Nuclear Safety, IEEE.

Seaver, D.A., and Stillwell, W.G.(1983), Procedures for Using Expert Judgement to Estimate Human Error Probabilities in Nuclear Power Plant Operations, Decision Science Consortium and Sandia National Laboratories, NUREG/CR-2743, US Nuclear Regulatory Commission, DC, March 1983.

Sebo, D.S., Dixon, B.W., and Bray, M.A.(1985), "Reactor Safety Assessment System (RSAS)," ANS Topical Meeting on Computer Applications for Nuclear Power Plant Operation and Control, Tri-Cities (Pasco), Washington, 8-12 September 1985, pp.721-727.

Shapero, A., Cooper, J.I., Rappaport, M., Schaeffer, K.H., and Bates, C.J.(1960), Human Engineering Testing and Malfunction data Collection in Weapon System Programs, Report WADD Technical Report 60-36, Wright Air Development Division , Wright -Patterson AFB, OH, February 1960. (cited in D.Meister, op. cit.).

Shepherd, A.(1976), "An Improved tabular format for task analysis, Journal of Occupational Psychology, 49, pp. 93-104.

Sheridan, T.B., Jenkins, J.P., and Kisner, R.A.(1982), Proceedings of Workshop on Cognitive Modelling of Nuclear Plant Control Room Operators, NUREG/CR-3114. Oak Ridge National Laboratory, Oak Ridge, Tennessee, 37830, USA.

Shortliffe, E.(1976), Computer-Based Medical Consultations: MYCIN, American Elsevier, New Work, NY.

Siegel, A.I., Bartter, W.D., Wolf, J.J., and Kneee, H.E.(1983), Front End Analysis for the Nuclear Power Plant Maintenance Personnel Reliability Model, Applied Psychological Swervices and Oak Ridge National Laboratory, NUREG?CR-2669, U.S. Nuclear Regulatory Commission, Washington, DC.

Sizewell B (1982) - A Review by H.M. Nuclear Installations Inspectorate of the pre-construction safety report, HMSO, London.

Stillwell, W.G., Seaver, D.A., and Swartz, J.P.(1982), Expert Estimation of Human Error Probabilities in Nuclear Power Plant Operations: A Review of Probablity and Scaling, NUREG/CR-2255, USNRC, May 1982.

Suokas, J, Heino, P., and Karvonen, I.(1987), "The Development of an Expert System to support HAZOP Analysis," Conference Proceedings of Reliability'87, Vols.1-2, April 1987, pp. 5B/R1/1-5B/R1/10.

Swain, A.D.(1963), A Method for Performing a Human factors Reliability Analysis, Monograph SCR-685, Sandia National Laboratories, Albuquerque, NM, August 1964.

Swain, A.D.(1963), A method for performing a human factors reliability analysis, Report SCR-685, Sandia National Laboratories, Albuquerque, NM, August 1963.

Swain, A.D.(1964), "Some Problems in the Measurement of Human Performance in Man-Machine Systems," Human Factors, vol.6, pp. 687-700.

Swain, A.D.(1987), "A Shortened Version of the THERP/Handbook Approach to Human Reliability Analysis for Probabilistic Risk Assessment," In: Proceedings of the Ergonomics Society's 1987 Annual Conference, Swansea, Wales, 6-10 April 1987, pp. 163-164.

Swain, A.D., and Guttmann, H.E., (1980-83), Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Applications, Sandia National Laboratories, NUREG/CR-1278, U.S. Nuclear Regulatory Commission, Washington, DC, October 1980 and August 1983.

Tanguy, P.(1988), "Three decades of nuclear safety," IAEA Bulletin, No. 2, pp.51-57.

Taylor, N.K., Corlett, E.N., and Simpson, M.R., "Problems of Knowledge Acquisition for Expert Systems," Proceeding of the Ergonomics Society's 1987 Annual Conference, Swansea, Wales, 6-10 April 1987, pp. 263-268.

Tharmalingam, S.(1989), "Assessing runaway reactions ans sizing vents," In: The Chemical Engineer, August 1989, pp33-41.

Thompson, J.R.(1987), Engineering Safety Assessment - An Introduction, Longman Science & Technical, UK, England.

Topmiller, D.A , Eckel, J.S., and Kozinsky, E.J (1983), "Review and Evaluation of Human Reliability Data Banks," pp. 541-549 In Proceedings of the International Meeting on THermal Nuclear Reactor Safety, 29 August 2 September, Chicago, IL, American Nuclear Society, NUREG/Cp-0027, U.S. Nuclear Regulatory Commission, Washington, DC, January 1983.

Topmiller, D.A., Eckel, J.S., and Kozinsky, E.J (1982), Human Reliability Data Bank for Nuclear Power Plant Operations, Volume 1: A Review of Existing Human Reliability Data Banks, General Physics Corporation and Sandia National Laboratories, NUREG/CR-2744, U.S. Nuclear Regulatory Commission, Washington, DC, December 1982.

Touchton, R.A., Gunter, D.A., and Cain, D.(1985), "Rule-Based Emergency Action Level Monitor Prototype," , Transactions of the American Nuclear Society, Winter Meeting, S.F., California, 10-14 November, 1985.

Trappl, R.(1986), ed., Impact of Artificial Intelligence, North Holland, Amsterdam.

U.K. Safety and Reliability Directorate (1981), Canvey: A Second Report, A Review of Potential Hazards from Operations in the Canvey Island, Thurrock Area, England, September 1981.

UKAEA (1988), Human Reliability Assessors Guide, Human Factors in Reliability Group, NCSR, UKAEA, Wigshaw Lane, Culcheth, Warrington.

US DoD (1977), Human Reliability System Users' Manual, USNAVSEA, Washington, DC, December 1977.

Waterman, D.A.(1985), Understanding Artificial Intelligence, Addison-Wesley, Reading, Mass.

Waterman, D.A.(1986), A Guide to Expert Systems, Addison-Wesley, Reading, Mass.

Watson, I.A.(1985), "Review of Human Factors in Reliability and Risk Assessment," In: IChemE Symposium Series No.93, pp. 323-351.

Weiss, S.M., and Kulikowsky, C.A.(1984), A Practical Guide to Designing Expert Systems, Rowman and Allanheld.

Whitfield, D. (1987a), "Human Reliability from a Nuclear Regulatory Viewpoint," Proceedings of the Ergonomics Society's 1987 Annual Conference, Swansea, Wales, 6-10 April 1987, pp. 58-63.

Whitfield, D. (1987b), "A Regulatory Perspective on Human Factors in Nuclear Power", Proceedings of a Conference on Human Reliability in Nuclear Power, IBC, London, 22nd-23rd October 1987.

Williams, H.L.(1958), "Reliability evaluation of the human component in man-machine systems", Electrical Engineering, April 1958, pp.78-82.

Williams, J.C.(1983), "Cost Effectiveness of Operator Features in Equipment Design," Applied Ergonomics, 14.2, pp.103-107.

Williams, J.C.(1986), "HEART - A Proposed Method for Assessing and reducing Human Error," In: Proceedings of the 9th Advances in Reliability Technology Symposium, University of Bradford, 4 April.

Woods, D.D., and Roth, E.M. (1986), "Modelling Cognitive Behaviour in Nuclear Power Plants: An Overview of Contributing Theoritical Traditions." In: ANS/ENS Proceedings of the International Topical Meeting on Advances in Human Factors in Nuclear Power Systems, Knoxvilee, Tennessee, 21-24 April 1986, pp. 12-20.

Zemva, D., Medjedovic, S., Piskar, R.(1987), "PC Based Expert System used to for MTBF and FMECA Analysis of Computer System," Conference Proceedings of Reliability'87, Vols.1-2, April 1987, pp. 5B/5/1 - 5B/5/6.

APPENDIX 1

ABBREVIATIONS

Following is a listing of the abbreviations used in the thesis.

ACMH	Advisory Committee on Major Hazards.
AHERAX	Absolute Human Error/Reliability eXpert.
AI	Artificial Intelligence.
APJ	Absolute Probability Judgement.
CIMAH	Control of Industrial Major Accidents Hazard Regulations.
DB	Data Bank/Base.
DOS	Disk Operating System.
DS	Data Store.
EEC	European Economic Commission.
EF	Error Factor.
ES	Expert System.
ET	Event Tree.
FT	Fault Tree.
GCLISP	Golden Common Lisp.
GEMS	Generic Error-Modelling System.
HASAWA	Health And Safety At Work Act.
HAZOP	HAZard and OPerability analysis.
HCR	Human Cognitive Reliability.
HEART	Human Error Assessment and Reduction Technique.
HEP	Human Error Probability.
HERAX	Human Error/Reliability Amalysis eXpert system.
HFs	Human Factors.

HMNII	Her Majest's Nuclear Installations Inspectorate
HMSO	Her Majesty's Stationary Office.
HR	Human Reliabilty.
HRA	Human Reliability Analysis.
HSC	Health and Safety Commission.
HSE	Health and safety Executive.
HTA	Hierarchical Task Analysis.
IAEA	International Atomic Energy Agency
IBM	International Business Management.
LISP	LISt Processing.
M-MI	Man-Machine Interface.
M-MS	Man-Machine System.
MAUD	Multi-Attribute Utility Decision.
MODEX-1/2	MODelling EXpert.
NII	Nuclear Industry Inspectorate.
NIIH	Notification of Installations Handling Hazardous substances.
NPP	Nuclear Power Plant.
OAT	Operator Action Tree.
PC	Paired Comparison.
PRA	Probabilistic Risk/Reliability Assessement/Analysis.
PSA	Probabilistic Safety Assessment/Analysis.
PSFs	Performance Shaping Factors.
QRA	Quantitative Risk/Reliability Assessment/Analysis.
QUANTEX	Quantification EXpert.
RFs	Recovery Factors.
SHARP	Systematic Human Action Reliability Procedure.
SHERAX	Subjective Human Error/Reliability Analysis eXpert system.
SLI	Success Likelihood Index.
SLIM	Success Likelihood Index Method.

STAHR	Socio-Technical Approach to Human Reliability.
TESEO	Technica Empirica Stima Errori Operatori.
THERAX	Technique for Human Error/Reliability eXpert system.
THERP	Technique for Human Error Rate Prediction.
TMI	Three Miles Island.
TRC	Time Correlation Curve.
UKAEA	United Kingdom Atomic Energy Authority.

APPENDIX 2

HERAX'S MAIN COMMANDS

The main commands used by the expert system HERAX are shown in the Table below:

COMMAND	USE
1. **WHY**	To get explanation why a specific question was asked.
2. **HOW**	To get explanation how a specific conclusion was reached.
3. **HELP**	To get help in answering a specific question.
4. **CTRL-BREAK**	To end the session and exit to DOS.
5. **YES/NO**	A Yes or No answer to a specific question is required.
6. **SELECT A NUMBER**	Type in the number of the specific question selected.
7. **ANY KEY TO CONTINUE**	Press any key to get the next screen display and continue the session.