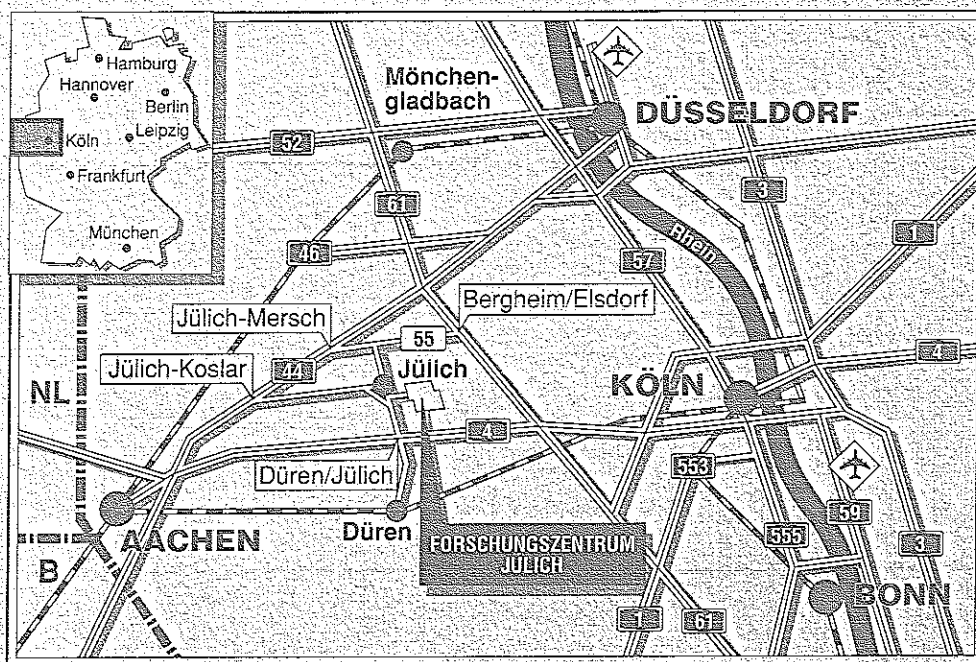


*Institut für Sicherheitsforschung
und Reaktortechnik*

**Development of Advanced Methods
and Related Software for Human
Reliability Evaluation within
Probabilistic Safety Analyses**

Kazimierz T. Kosmowski
Johannes Mertens

Gisela Degen
Bernhard Reer



Berichte des Forschungszentrums Jülich ; 2928

ISSN 0944-2952

Institut für Sicherheitsforschung und Reaktortechnik Jüli-2928

Zu beziehen durch: Forschungszentrum Jülich GmbH · Zentralbibliothek

D-52425 Jülich · Bundesrepublik Deutschland

Telefon: 024 61 / 61 - 61 02 · Telefax: 024 61 / 61 - 61 03 · Telex: 833 556-70 kfa d

Development of Advanced Methods and Related Software for Human Reliability Evaluation within Probabilistic Safety Analyses

Kazimierz T. Kosmowski*
Johannes Mertens

Gisela Degen
Bernhard Reer

* Technical University, Gdansk

**Development of Advanced Methods
and Related Software
for Human Reliability Evaluation
within Probabilistic Safety Analyses**

Kazimierz T. Kosmowski*

Gisela Degen

Johannes Mertens

Bernhard Reer

Abstract

Human Reliability Analysis (HRA) is an important part of Probabilistic Safety Analysis (PSA). The first part of this report consists of an overview of types of human behaviour and human error including the effect of significant performance shaping factors on human reliability.

Particularly with regard to safety assessments for nuclear power plants a lot of HRA methods have been developed. The most important of these methods are presented and discussed in the report, together with techniques for incorporating HRA into PSA and with models of operator cognitive behaviour.

Based on existing HRA methods the concept of a software system is described. For the development of this system the utilization of modern programming tools is proposed; the essential goal is the effective application of HRA methods. A possible integration of computer-aided HRA within PSA is discussed.

The features of Expert System Technology and examples of applications (PSA, HRA) are presented in four appendices.

*Technical University, Gdansk

**Entwicklung fortschrittlicher Methoden
und zugehöriger Computercodes
zur Bewertung menschlicher Zuverlässigkeit
im Rahmen probabilistischer Sicherheitsanalysen**

Kazimierz T. Kosmowski*

Gisela Degen

Johannes Mertens

Bernhard Reer

Kurzfassung

Die Untersuchung menschlicher Zuverlässigkeit (Human Reliability Analysis, HRA) ist ein bedeutsamer Teil probabilistischer Sicherheitsanalysen (PSA). Der erste Teil dieses Berichtes besteht aus einem Überblick zu den Arten menschlichen Verhaltens und menschlicher Fehler einschließlich der signifikanten Einflußgrößen auf menschliche Zuverlässigkeit.

Insbesondere für Sicherheitsbewertungen von Kernkraftwerken wurde eine Reihe von HRA-Methoden entwickelt. Die wichtigsten von ihnen werden vorgestellt und diskutiert, zusammen mit Vorgehensweisen zur Einbeziehung von HRA in PSA sowie mit Modellierungen des kognitiven Verhaltens von Betriebsmannschaften.

Auf der Basis vorhandener HRA-Methoden wird das Konzept eines Computer-Programmsystems beschrieben. Zur Entwicklung dieses Systems wird die Verwendung moderner Programmiermethoden vorgeschlagen; das Hauptziel ist die wirkungsvolle Anwendung von HRA-Methoden. Eine mögliche Einbeziehung computerunterstützter HRA in PSA wird diskutiert.

Die grundsätzliche Besonderheit von Expertensystemen und Anwendungsbeispiele (PSA, HRA) werden in vier Anhängen vorgestellt.

*Technische Universität, Danzig

Contents

Contents	i
List of Figures	vii
List of Tables	ix
List of Abbreviations	xi
1 Introduction	1-1
2 Techniques for human reliability analysis	2-1
2.1 Introductory remarks	2-1
2.2 Classification of human actions and errors	2-2
2.2.1 Human behaviour types	2-2
2.2.2 Error types	2-3
2.2.3 Incorrect human outputs	2-4
2.2.4 Phase of actions	2-4
2.2.5 Performance Shaping Factors (PSFs)	2-5
2.3 HRA techniques	2-7
2.3.1 General remarks	2-7
2.3.2 Time-dependent methods	2-8
2.3.2.1 Time Reliability Correlation (TRC)	2-8
2.3.2.2 Human Cognitive Reliability (HCR) model	2-10
2.3.2.3 Operator Action Tree (OAT)	2-12
2.3.3 Decomposition methods	2-14
2.3.3.1 Technique for Human Error Rate Prediction (THERP)	2-14
2.3.3.2 Accident Sequence Evaluation Procedure, Human Reliability Analysis Procedure (ASEP-HRAP)	2-16
2.3.3.3 Human Error Assessment and Reduction Technique (HEART)	2-17
2.3.3.4 INTENT	2-18
2.3.4 Expert estimation methods	2-19
2.3.4.1 Success Likelihood Index Method (SLIM)	2-19
2.3.4.2 Socio-Technical Approach to Assessing Human Reliability (STahr)	2-20
2.3.4.3 Pair Comparison (PC) technique	2-22
2.3.4.4 Absolute Probability Judgement (APJ)	2-23

2.3.4.5 Confusion Matrix (CM) technique	2-26
2.3.5 Expert estimation methods seldom applied in PSA studies and other conventional HRA methods	2-28
2.3.5.1 MAintenance Personnel Performance Simulation (MAPPS) model	2-28
2.3.5.2 TESEO	2-29
2.3.5.3 Other conventional HRA methods	2-29
2.3.6 An example of software supporting HRA	2-29
2.4 Frameworks for incorporating HRA into PSA	2-32
2.4.1 Systematic Human Action Reliability Procedure (SHARP)	2-32
2.4.2 Procedure for Conducting Human Reliability Analysis in Probabilistic Safety Assessment (IAEA's Procedure)	2-34
2.5 Approaches based on integrated simulation or AI methods	2-37
2.5.1 Cognitive Environment Simulation (CES)	2-37
2.5.2 An integrated modelling of operator cognitive behaviour	2-40
2.5.3 COgnitive SIMulation MOdel (COSIMO)	2-44
2.6 Remarks on available HRA techniques	2-48
2.6.1 Evaluation of quantitative HRA methods	2-48
2.6.1.1 Human Cognitive Reliability (HCR) model	2-48
2.6.1.2 Technique for Human Error Rate Prediction (THERP)	2-48
2.6.1.3 Human Error Assessment and Reduction technique (HEART)	2-48
2.6.1.4 INTENT	2-49
2.6.1.5 Success Likelihood Index Method (SLIM)	2-49
2.6.1.6 Socio-Technical Approach to Assessing Human Reliability (STahr)	2-49
2.6.1.7 Paired Comparison (PC) technique	2-49
2.6.1.8 Absolute Probability Judgement (APJ)	2-50
2.6.1.9 Confusion Matrix (CM) technique	2-50
2.6.1.10 Summary	2-50
2.6.2 Evaluation of advanced (qualitative) HRA methods	2-51
3 Concept of the software system supporting human reliability evaluation and documentation	3-1
3.1 Methodological issues of HRA	3-1
3.1.1 Human interactions and assumptions concerning HRA/PSA	3-1
3.1.2 Types of human behaviour and errors	3-2
3.1.3 Recovery potential	3-5
3.2 Classification of human actions/errors and quantification techniques	3-7

3.2.1 The HRA procedure	3-7
3.2.2 Human actions/errors and selection of quantification techniques	3-11
3.2.3 HRA/PSA effort and selection of quantification techniques	3-13
3.3 Quantitative analysis of human reliability	3-13
3.3.1 General remarks	3-13
3.3.2 Quantifying slips	3-14
3.3.3 Quantifying mistakes	3-17
3.4 Proposed method for the quantification of decision-based errors in HRA	3-17
3.4.1 State of research	3-18
3.4.2 Description of the SLIM-based model	3-21
3.4.3 Description of the PSF's selected to operationalize risk-taking behaviour	3-23
3.4.4 Using <i>one</i> reference HEP for calibration	3-24
3.4.5 Uncertainties	3-24
3.4.6 Related software	3-25
3.4.7 Future research fields	3-26
3.5 Software supporting human reliability analysis and evaluation	3-26
3.5.1 Concept of software system	3-27
3.5.2 An expert system for the application of THERP	3-28
3.6 Main steps for further system development	4-1
4 Integration of computer-aided HRA with PSA	4-1
4.1 Incorporation of human-induced events into PSA logic	4-4
4.2 Application of the ES technology for HRA/PSA	4-8
4.3 Hardware and software environment for advanced HRA/PSA	4-9
4.4 Methodological issues for treating uncertainty in HRA/PSA	4-9
4.4.1 Bayesian and non-Bayesian theories	4-11
4.4.2 Combining knowledge	4-14
4.4.3 Representing uncertainties and combining information in expert systems	4-15
4.4.4 Treatment of uncertainties in computerized PSA/HRA	4-19
4.5 Proposals for system development	5-1
5 Conclusions	R-1
References	

Appendix 1

Features of the expert system technology

- A1.1 Expert Systems and Decision Support Systems
- A1.2 Software for building engineering expert systems

A1-1

A1-1

A1-5

Appendix 2

A hybrid software tool for building expert systems:

Knowledge Engineering Environment (KEE)

- A2.1 General remarks
- A2.2 The structure of KEE
 - A2.2.1 Units
 - A2.2.2 Slots
 - A2.2.3 Facets
 - A2.2.4 Methods
 - A2.2.5 Active values
 - A2.2.6 Rules
 - A2.2.6.1 Forward chaining
 - A2.2.6.2 Backward chaining

A2-1

A2-1

A2-2

A2-2

A2-3

A2-4

A2-4

A2-5

A2-6

A2-7

A2-9

Appendix 3

An expert system for the application of THERP

- A3.1 Introduction
- A3.2 Content conversion of THERP
 - A3.2.1 Procedure
 - A3.2.2 Calculations
 - A3.2.2.1 General
 - A3.2.2.2 Basic failure probabilities
 - A3.2.2.3 Modified failure probabilities
 - A3.2.2.4 Recovery failure probabilities and failure conditions
 - A3.2.2.5 Error type contributions and step failure probability
 - A3.2.2.6 Failure path probabilities
 - A3.2.2.7 Procedure failure probabilities and system failure probability
 - A3.2.2.8 Summary
- A3.3 Programme conversion of THERP into KEE
 - A3.3.0 General
 - A3.3.1 Modelling the structure

A3-1

A3-1

A3-2

A3-2

A3-5

A3-5

A3-7

A3-7

A3-9

A3-9

A3-10

A3-11

A3-11

A3-12

A3-12

A3-12

A3.3.1.1 Units	A3-12
A3.3.1.2 Slots	A3-14
A3.3.2 Inputs	A3-16
A3.3.2.1 Entering the hierarchy	A3-16
A3.3.2.2 Data input for procedures	A3-18
A3.3.2.3 Data for steps	A3-19
A3.3.3 Calculation of the probability and output of results	A3-22

Appendix 4

Examples of expert systems supporting Probabilistic Safety Analyses	A4-1
A4.1 LESSEPS	A4-1
A4.2 PRA-EXPERT and REX	A4-4
A4.3 SAPIEXT	A4-6
A4.4 STARS	A4-7
A4.5 Other expert systems	A4-9

List of Figures

Fig. 2.1	Illustration of Rasmussen's conceptual framework of three cognitive levels of human behaviour /Ra83/	2-2
Fig. 2.2	The basic time reliability correlation /Do88/	2-9
Fig. 2.3	Normalized crew non-response curves for skill-, rule-, and knowledge-based behaviour /Hu88/	2-11
Fig. 2.4	Expanded Operator Action Tree (EOAT) representation /Hu88/	2-13
Fig. 2.5	Human Reliability Analysis Event Tree (HRAET) for series and parallel systems	2-15
Fig. 2.6	The STAHR influence diagram /Ph89/	2-21
Fig. 2.7	Example of confusion matrix (CM) for misdiagnosis of an event /IA91c/	2-27
Fig. 2.8	ORCA logic structure /Do88/	2-31
Fig. 2.9	Links between SHARP steps /Ha84/	2-32
Fig. 2.10	Different kinds of processing in CES /Wo87/	2-38
Fig. 2.11	Cognitive Reliability Assessment Technique (CREATE) /Wo87/	2-39
Fig. 2.12	General framework of the human behaviour model /Am87/	2-42
Fig. 2.13	Goal, task and act concept /Am87/	2-43
Fig. 2.14	Conceptual architecture of the COSIMO model /Ca89/	2-46
Fig. 2.15	Blackboard basic cycle /Ca89/	2-46
Fig. 3.1	Rasmussen's conceptual framework of human behaviour and decision making	3-2
Fig. 3.2	Dynamics of Generic Error-Modelling Systems (GEMS) /Re87/	3-3
Fig. 3.3	The HRA process /Do88/	3-8
Fig. 3.4	Classification proposal for human pre-incident actions/errors and quantification technique	3-9
Fig. 3.5	Classification proposal for human event-driven action/errors and quantification techniques	3-10
Fig. 3.6	Methods to assess the probability of decision-based errors in hazardous situations	3-21
Fig. 4.1	Levels of incorporation of human failure events in PRA models /Do88/	4-2
Fig. 4.2	Recovery fault tree model for recirculation failure event E /Do88/	4-3
Fig. 4.3	Functions and processing phases of a PSA/HRA expert system	4-5
Fig. A1.1	Structure of an expert system /Tu90/	A1-1

Fig. A1.2	Expert system development process /Tu90/	A1-2
Fig. A1.3	Stages of knowledge acquisition /Tu90/	A1-3
Fig. A1.4	Software for building expert systems /Tu90/	A1-4
Fig. A1.5	Expert System as a separate component of Decision Support System /Tu90/	A1-4
Fig. A2.1	Example of a unit hierarchy with terms from reactor technology	A2-2
Fig. A2.2	Unit hierarchy of energy sources	A2-4
Fig. A2.3	Unit hierarchy of bodies	A2-5
Fig. A2.4	NPP system in a unit hierarchy	A2-7
Fig. A2.5	Unit hierarchy of simulation rules	A2-7
Fig. A3.1	Simplified fault tree for FRJ-2 emergency cooling	A3-2
Fig. A3.2	Computation rules for lognormal-distributed probabilities	A3-6
Fig. A3.3	Hierarchy for a system failure	A3-12
Fig. A3.4	Subdivision of unit class "Procedures"	A3-13
Fig. A3.5	Classification	A3-13
Fig. A3.6	Unit hierarchy of THERP in KEE	A3-14
Fig. A3.7	Hierarchy prior to the 1st program start	A3-14
Fig. A3.8	Menu	A3-17
Fig. A3.9	Data input for a procedure prior to the accident	A3-18
Fig. A3.10	Entering the sequences	A3-20
Fig. A3.11	Entering the direct steps	A3-20
Fig. A3.12	Entering the two combinations	A3-21
Fig. A3.13	Entering the important failure paths	A3-22
Fig. A3.14	Rule hierarchy	A3-23
Fig. A3.15	Rule hierarchy for W_1	A3-24
Fig. A4.1	General principles of EXPRESS and EXPGSI expert systems /Ma90/	A4-1
Fig. A4.2	Express and EXPGSI processing sequences /Ma90/	A4-2
Fig. A4.3	A simple Master Plant Logic Diagram (MPLG) /Wa90/	A4-4
Fig. A4.4	Overview of tasks covered in STARS /Po90/	A4-8

List of Tables

Tab. 2.1	Rasmussen's behaviour type classification	2-2
Tab. 2.2	The slip/mistake dichotomy /Re87, Do88/	2-3
Tab. 2.3	Human actions/errors related to the phases of an accident /Do88, IA92b/	2-5
Tab. 2.4	Classification of actions/errors in event-driven activities /IA92a, IA92b/	2-5
Tab. 2.5	Methods for obtaining evaluation from experts	2-25
Tab. 2.6	Type of human interactions according to SHARP	2-32
Tab. 2.7	SHARP steps	2-33
Tab. 2.8	Main components of an HRA study /IA92b/	2-35
Tab. 2.9	Incorporating HRA into PSA	2-36
Tab. 2.10	Quantitative HRA methods and criteria used for evaluation	2-51
Tab. 3.1	Contributions of intrinsic and extrinsic factors in shaping specific error forms /Re87/	3-4
Tab. 3.2	A summary of the psychological varieties of unsafe acts /Re90a, p.207/	3-17
Tab. 3.3	Complicating factors - some variations or difficulties that go beyond the standard method for handling the situation /Wo90b, p.44/	3-18
Tab. 3.4	Selected PSF's for the probabilistic quantification of risk taking behaviour	3-22
Tab. 3.5	Discretization of PSF ratings	3-23
Tab. 3.6	Future research fields concerning a model for the probabilistic quantification of risk taking behaviour	3-25
Tab. 4.1	Assumed features of software supporting PSA-level-1 and HRA based on the expert system technology	4-6
Tab. A2.1	Slots defined in the hierarchy of Fig. A2.1	A2-3
Tab. A2.2	Slot values before rule application in forward chaining example	A2-8
Tab. A2.3	Slot values after rule application in forward/backward chaining example	A2-9
Tab. A2.4	Slot values before rule application in backward chaining example	A2-9
Tab. A3.1	Questions concerning the procedure prior to an accident	A3-3
Tab. A3.2	Questions for perception errors	A3-4
Tab. A3.3	Continuation of Tab. A3.2	A3-5
Tab. A3.4	Continuation of Tab. A3.2	A3-5
Tab. A3.5	Median (uncertainty factor) of the basic failure probability of U_1	A3-7
Tab. A3.6	Median (uncertainty factor) of the basic failure probability of W_1	A3-8

Tab. A3.7	Modification of the median of the basic failure probability of U_1	A3-8
Tab. A3.8	Median (uncertainty factor) of a recovery failure probability for procedures prior to an accident	A3-9
Tab. A3.9	Median of the conditional step failure probability $SF_{i/i-1}$	A3-11
Tab. A3.10	Dynamic data slot changes	A3-16
Tab. A3.11	Conditions for input	A3-17
Tab. A3.12	Conditions for step data input	A3-19
Tab. A3.13	Main menu	A3-25

List of Abbreviations

AI	Artificial Intelligence
ANS	American Nuclear Society
APJ	Absolute Probability Judgement /Se83/
ASEA	Assistant Sequence Evaluation Procedure /Sw87/
CAD	Computer-Aided Design
CCDF	Complementary Cumulative Distribution Function
CES	Cognitive Environment Simulation /Wo86, Wo87/
COSIMO	Cognitive Simulation Model /Ca89/
CM	Confusion Matrix /Po81/
CREATE	Cognitive Reliability Assessment TEchnique /Wo87/
DYLAM	Dynamic Logical Analytical Methodology /Am87, Ca88/
DSS	Decision Support System
D&D	Diagnosis or Decision Making
ECOM	Error of Commission
EF	Error Factor
ENS	European Nuclear Society
EOI	Error of Intention
EOM	Error of Omission
EPC	Error Producing Condition (related to HEART)
ESAP	Expertensystem zur systematischen Analyse von Personalhandlungen in Risikostudien /Re93/
GEMS	Generic Error Modelling System /Re87/
HCR	Human Cognitive Reliability /Ha85b/
HEART	Human Error Assessment and Reduction Technique
HEP	Human Error Probability
HLDM	High Level Decision Making
HRA	Human Reliability Analysis
HRAET	Human Reliability Analysis Event Tree
HRAP	Human Reliability Analysis Procedure /Sw87/
INTENT	see /Ge92/
KB	Knowledge Base(s)
KEE	Knowledge Engineering Environment
LRDM	Low Level Decision Making
LER	Licensee Event Report
MAPPS	Maintenance Personal Performance Simulation /Si84/
MAUD	Multi-Attribute Utility Decomposition
MAUT	Multi-Attribute Utility Theory
MCS	Minimal Cut Set

NPP	Nuclear Power Plant
NUCLARR	Nuclear Computerized Library for Assessing Reactor Reliability /Ge89/
OAT	Operator Action Tree /Wr82, Ha82/
ORCA	Operator Reliability Calculation and Assessment /Do88/
PAF	Performance Adjustment Factor (related to CES)
PC	Paired Comparison /Se83/
PFS	Performance Shaping Factor
PRA	Probabilistic Risk Assessment
pr(...)	Probability that...?
PSA	Probabilistic Safety Assessment
QA	Quality Assurance
SARAH	Systematic Approach to the Reliability Assessment of Humans
SHARP	Systematic Human Action Reliability Procedure /Ha84/
SLIM	Success Likelihood Index Methodology /Em84/
STAHR	Socio-Technical Approach to Assessing Human Reliability /Ph89/
TESEO	Tecnica Empirica Stima Errori Operatori /Be80/
THERP	Technique of Human Error Rate Prediction /Sw83/
TRC	Time Reliability Correlation(s)

1 Introduction

Probabilistic Safety Analysis (PSA) has been recognized as a useful tool supporting decision making during the design, licensing and operation of complex hazardous technological systems. PSA studies have already been performed for more than 15 years. However, lately critical remarks have been raised within the PSA research and engineering community concerning methodologies currently practically applied. These are associated with treating the human factor and logical/probabilistic modelling of safety systems, the scope of modelling dependent failures, eliciting and aggregating of expert opinions, quantifying the frequency of rare events and events not supported by statistical data at all, etc. One of the most important and difficult issues is reliability modelling of the so called human factor.

Results of PSA benchmark exercises, e.g. that summarized in /Po88a, Po88b/, have shown that the assessed frequency of some accident conditions for given power plants, obtained by several research groups, can display a discrepancy as high as several orders of magnitude. There is, therefore, an understandable trend at some research institutions to improve methodology and to propose a way to standardize human reliability analysis (HRA) and PSA studies. PSA is of a prime importance from the point of view of usefulness of quantitative results for reliability management and safety related decision making for hazardous systems. In order to accomplish these objectives growing interest in a wide computerization of PSA studies and, in particular, applying the Expert Systems Technology (EST) can be observed.

Computer-based technologies are now being intensively developed to improve the effectiveness of managerial decision making, especially in complex tasks. Three technologies can be distinguished: Decision Support Systems (DSS), Executive Information Systems (EIS) and Expert Systems (ES). They are collectively known as computerized management support systems. Expert systems, which appeared commercially in the early 1980s, may be considered as one of the most important future breakthroughs in computerized decision making /Tu90/.

Expert Systems (ESs) that emerged from Artificial Intelligence (AI) methods and object-oriented programming are designed to possess a feature to simulate expert reasoning on the knowledge base developed under the supervision from the experts from the given domain. They are especially useful to solve complex problems for which formal methods are rather limited and heuristics is significantly involved. Experts performing PSA studies face similar complex problems, often difficult to describe using formal mathematical methods. EST enables these studies to be facilitated and documented. Repeatable results will be obtained for given data/knowledge bases and defined reasoning strategy. The contribution of experts will be documented and it will thus be possible to scrutinize and audit data and results.

It has been already recognized that the introduction of advanced computer technology is a prerequisite for satisfactorily managing and improving the reliability and safety of various modern technological systems. There has been growing interest lately in applying expert systems in many domains. ESs are being used in a wide variety of applications including diagnosis, planning, forecasting, design, interpretation, control, status monitoring, instruction, etc. The industrial sector has identified expert systems as a means of: increasing plant safety; distributing the knowledge of experts to less qualified workers; increasing industrial reliability, throughput and economy; and preserving the expertise of retired workers.

This report is devoted to methods and related software for human reliability evaluation. The objective is to propose the development of computerized tools for supporting Human Reliability Analysis (HRA) performed within PSA studies of various complex technological systems, nuclear power plants in particular. These tools should be developed in such a way as to facilitate and standardize relevant studies. It was decided to also discuss some methodological issues of HRA/PSA in this report, e.g. the issue of uncertainty representation and treatment. The report consists of an outline of the future research and design work to develop and implement the computerized HRA/PSA tools.

Designing as well as verification and validation of expert systems requires a considerable effort, so that a step by step procedure of their development is proposed. The consequent expert systems should be developed in such a way as to enable their future integration into a complex and advanced expert system for semi-automatic PSA for various hazardous technological systems.

This report consists of five sections, references and four appendixes. In section 2 some methodological issues of HRA are considered and several HRA methods are described and evaluated. In section 3 the concept of the software system supporting human reliability evaluation is outlined based upon human cognitive behaviour theories. In section 4 a way of software system integration for computer-aided HRA with the software for computer-supported PSA is proposed and the problem of treating uncertainty is discussed.

2 Techniques for human reliability analysis

2.1 Introductory remarks

The influence of the so called human factor on the reliability and safety of complex industrial systems, Nuclear Power Plants (NPPs) in particular, is widely recognized. Due to severe accidents that have occurred in several industrial and transportation systems, e.g. TMI, Chernobyl, the Bhopal chemical plant and several refineries, the Challenger shuttle, public opinion and the scientific community are more and more convinced that the contribution of the human factor to the risk is significant.

After these accidents more research work has been undertaken focused on key issues of the man-machine interface to minimize the possibility of human errors. Human errors can be committed during the design, construction and operation of the plant. This report is devoted to the analyses of human errors committed during plant operation, although results of the relevant studies can be useful for its design. Because the human factor interacts with the system, the analyses should be useful for comparing the contribution of both man and machine to the system performance, thus it is necessary to find comparable measures. The probabilistic measures proved to be convenient for a description of the system performance as regards its reliability and safety. The role of Human Reliability Analysis (HRA) is more significant for complex dynamic process systems.

Several more important techniques are available for performing HRA. First human reliability models were developed based upon behavioural theories. In such models man was treated like a machine. Recently research interest has been focused on cognitive psychology but only a few new models have been proposed. They are not fully mature and require validation. Several research programmes have been carried out on simulators to confirm the HRA models. The effects of this research are in some cases controversial. There has recently been a call for the development of a "third generation" of human reliability models /Do90/. Significant research efforts have been undertaken in some research institutions to apply Artificial Intelligence (AI) methods combined with cognitive psychology theories.

This report addresses some important issues of HRA and practical aspects of the integration of HRA within Probabilistic Safety Analysis (PSA). Several conventional HRA techniques have been found to be appropriate for the evaluation of Human Error Probability (HEP). The selected techniques account for several types of the error mechanisms. Some methods are suited to investigate the influence of Performance Shaping Factors (PSFs) to interpolate estimates from other methods. Because the HRA and PSA studies are complex it is proposed to develop special software to support experts in quantitative evaluations and reliability-related decision making.

2.2. Classification of human actions and errors

2.2.1. Human behaviour types

The distinction between three categories of human behaviour has been proposed by Rasmussen /Ra83,Ra87/. His conceptual framework (Fig. 2.1) assumes three cognitive levels of human behaviour:

I	skill-based	frequently practised tasks that can be performed as more or less subconscious routines governed by stored patterns of behaviour
II	rule-based	performance of less familiar tasks in which a person follows remembered or written rules
III	knowledge-based	performance of novel actions where familiar patterns and rules cannot be applied directly, and actions follow the information processing with inclusion of planning and decision making

Tab. 2.1: Rasmussen's behaviour type classification

Skill-based tasks require little or no decision making and hence the task goes from initial stimuli from sensory inputs, i.e. "activation", to the realization of sensori-motor patterns, i.e. "execution" (see also Fig. 3.1). Rule-based tasks require some decision making to proceed from the initial stimuli and recognition to the "integration" phase where information is then processed an appropriate procedure or rule selected. Finally the selected procedure or rule is executed /IA92b/.

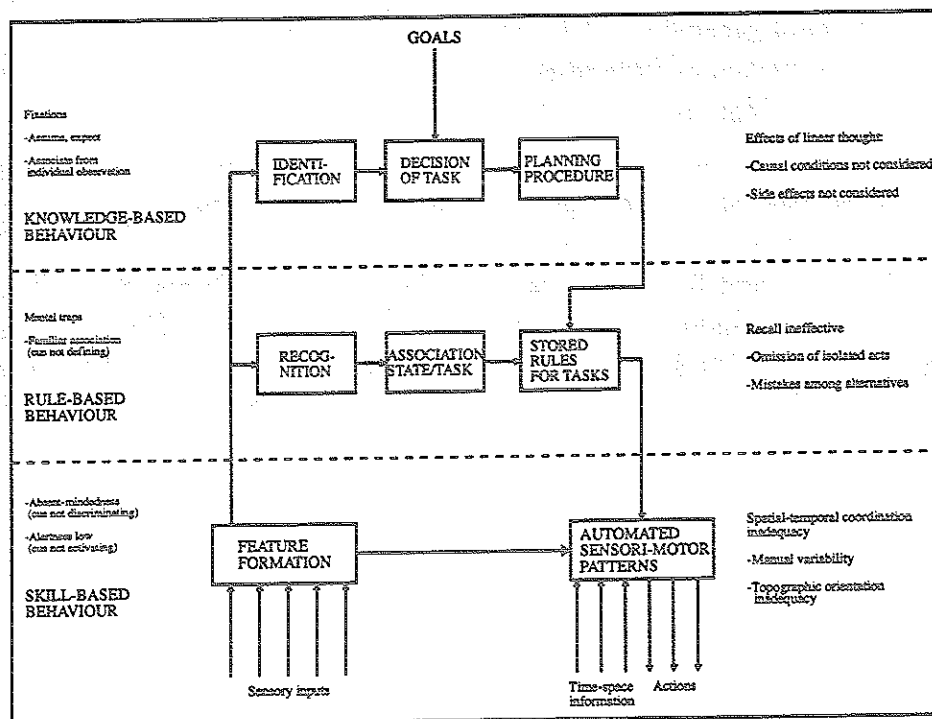


Fig. 2.1: Illustration of Rasmussen's conceptual framework of three cognitive levels of human behaviour /Ra83/

Knowledge-based behaviour involves cognitive activities such as identification, planning and a comparatively high degree of decision making. The relevant process of human behaviour includes the interpretation and evaluation of information, planning of actions and the execution of a plan. The plan can consist partly of some appropriate procedures or rules. Knowledge-based behaviour is often necessitated by uncertainty or novelty in the circumstances at hand.

It is known from HRA practice that the distinction between skill-based actions and rule-based actions is often more or less arbitrary. The similar difficulty is often to be faced with the distinction between rule-based and knowledge-based behaviour. This is confirmed by new psychological theories /Re87/.

2.2.2. Error types

The behaviour types described above seem to involve different error mechanisms, which may mean radically different reliability characteristics /IA92b/. Human errors are often classified to be of two kinds /Re87, Do88/:

I	slip	is (1) an error in implementing a plan, decision or intention (the plan is correct, its execution is not), or (2) an unintended action; a type of slip is a lapse, an error in recall, e.g. of a step in a task.
II	mistake	is an error in establishing a course of actions, e.g. an error in diagnosis, planning or decision making.

Tab. 2.2: The slip/mistake dichotomy /Re87, Do88/

Slips/lapses can occur during the execution of skill-based actions. Slip is e.g. inadvertent selection of a wrong item when attempting to execute a planned action. Lapses are omissions of executions of a planned sequence of actions.

Mistakes can be committed when the rule- and knowledge-based actions are to be planned. They are associated with more serious error mechanisms as they lead to incorrect understanding of a situation and conceiving an inappropriate plan of resulting actions or sequences of actions. Mistakes can also occur due to inappropriate selection of familiar procedures or rules based only on part of the information potentially available, in order to reduce the amount of decision making required /IA92b/.

Slips usually occur in a less time-constrained environment, it is, therefore, reasonable to believe that their associated error probabilities are less time-dependent, more structural, and arise as faults in an intended protocol or procedure. On the other hand, mistakes occur more often in a time-constrained environment and are more time-dependent (it takes "process" time to think up an appropriate response in a novel situation). Mistakes are less structural, and often have little to do with protocol or procedure /Do88/.

In reality, an overall task consists of "slip-likely" and "mistake-likely" subtasks. The performance of every subtask will take time. Therefore, time-dependency should be quantified in an overall approach /Re92a/.

For PSA purposes it is important to be aware of the different potential for error recovery for these types of error mechanisms. Slips and lapses can generally be recovered fairly quickly providing there are appropriate feedback mechanisms and plant behaviour is reversible. On the other hand, mistakes are usually less easily recovered in the short term. "Mindset" problems can occur and operators can persist in attempting to implement an inappropriate plan even when faced with a lot of contradictory information. Recovery routes in such situations have to be very positive and powerful to be reliable /IA92b/.

2.2.3. Incorrect human outputs

An early classification scheme for human errors, used in THERP /Sw83/, is related to the observable part of undesirable human behaviour, i.e. incorrect outputs. Human errors are categorized as either commissions or omissions. This classification assumed that the reliability of tasks could be assessed by decomposing them into subtasks much in the same way that routine task performance can be institutionalized by writing step-by-step procedures. If a step was left out, an error of omission was said to have occurred, and if a step was otherwise performed incorrectly, an error of commission occurred /Do88/.

Error of commission (ECOM) is incorrect performance of a system-required task or action, or the performance of some extraneous task or action that is not required by the system and which has the potential for contributing to some system-defined failure. Error of omission (EOM) is a failure to perform a task or action /Sw88/.

This scheme of error partition has no direct relevance to the mechanisms of their causes. For example, a person could just as easily commit an error as to omit a step because of, say, a lapse in attention. As a result, there would be no reason to expect a difference in any reliability parameter based only on this distinction /Do88/.

2.2.4. Phase of actions

In PSA two main phases are considered: the time prior to the initiator and the time after the initiator of an abnormal event. The role of personnel including operators changes over these time periods: from normal maintenance and operation activities, which are directed by familiar procedures, to incident investigation and management, and then, infrequently, to emergency response. These latter activities are driven by the event at hand and its perceived severity /Do88/.

Thus human interactions (actions/errors) can have a variety of means of affecting safety. It is important to be able to relate them to the PSA structure to understand their potential

effect on risk or safety. Human actions/errors can be divided into three categories so as to relate to the phases of an accident /Do88, IA92b/:

A	Actions/errors in planned activities, i.e., so-called pre-initiator events that cause equipment/systems to be unavailable when required post-initiator.
B	Errors in planned activities that lead directly, either by themselves or in combination with equipment failures, to initiating events/faults (e.g. unplanned reactor shutdown), i.e. human-induced initiators.
C	Actions/errors in event-driven (off-normal) activities, i.e., post-initiator events; these can be either performing safety actions or errors that aggravate the fault sequence.

Tab. 2.3: Human actions/errors related to the phases of an accident /Do88, IA92b/

Interactions of category C can be separated into three different types for incorporation into the PSA:

C1	Procedural safety actions. These actions involve success or failure in following procedures or rules known to operators in response to an accident sequence.
C2	Aggravating actions/errors. These interactions are a special set of commission errors that occur post-fault and significantly aggravate the accident progression. They are the most difficult to identify and model.
C3	Improvising recovery/repair actions. These consist of recovery actions, which are usually only included in accident sequences that dominate risk profiles. They may include the recovery of previously unavailable equipment or the use of non-standard procedures to ameliorate the accident conditions.

Tab. 2.4: Classification of actions/errors in event-driven activities /IA92a, IA92b/

Maintenance activities (i.e. maintenance, test, repair, calibration and inspection) can lead to category A or B events. Control actions/errors during normal operation can either lead to events of category B or A, e.g. failing to realign systems/equipment after testing, omitting to restore protection system vetoes or permissions. Events of category C can occur either in the control room or locally at the plant /IA92b/.

The first two categories (A and B) are behaviourally identical, differing only in whether the results of the failure immediately upset plant phenomena or not /Do88/. Actions of types C1, C2 and C3 are all involved in operator response to an accident once it has been initiated. Often, some diagnosis is required and time is a limiting factor. The general approach for dealing with type C1 and C3 interactions is the same /IA92b/.

2.2.5. Performance Shaping Factors (PSFs)

Any factor that influences the reliability of human performance is designated as a PSF. Many factors have been distinguished which can potentially influence the behaviour of human beings

as a part of the man-machine system. They were divided into three classes: (1) the external PSFs: those outside the individual; (2) the internal PSFs: those that can be activated within the individual himself; and (3) stressors: psychological and physiological /Sw83, IA91c/.

In HRA only a few PSFs are usually considered, depending on the models applied. One of the most important PSFs difficult to include in HRA models is stress. Other important PSFs are: quality of control room design, quality of training and other quality assurance aspects including administrative procedures, operational procedures and operator redundancy /Sw83, Iv92/.

It should be emphasized that modelling the effect of PSFs on human performance can be only approximate. In the present state of HRA it is not an easy task to assign weighting factors to each individual PSF that would be applicable to specific tasks and situations analysed. The problem is that in reality most PSFs interact, and the interactions are usually complex. The SLIM technique described in item 2.3 provides an analytical frame for combining the PSFs to determine the overall impact on a relative basis between different tasks. Another alternative is to consider only a fixed number of PSFs for any kind of task in a defined pattern of interaction, as proposed e.g. in the HCR model.

2.3 HRA techniques

2.3.1 General remarks

It would be desirable to develop the human reliability data base consisting of "hard" data through the evaluation of raw performance data from the industry and experiments carried out on laboratory stands or simulators. Such trials are known and some compiled data exist. The way from raw data to useful information is, however, laborious and biased with human judgement /Do88/.

It is known that in reliability engineering, raw data takes the form of failure records, opportunity counts, and attribute lists. Users of raw data typically select among the data according to some set of attributes and evaluate the failure statistics for that "population". The result is compiled data. Further judgement is involved to fit the data to distributions and determine statistical parameters. The result can then be called interpreted data /Do88/.

Such an approach is being applied for sets of hardware components for which the occurrence of failure events is comparatively frequent. In cases where the occurrence of a failure is a rare event the estimates of the frequency of occurrence can be made by using other techniques, e.g. estimates with regard to similar components with more frequent occurrence or aggregation of data for similar devices to increase the population of similar components. Such approaches also depends heavily on expert judgement /Do88/.

A similar approach in HRA is usually inappropriate, especially when events to be estimated are rare and aggregation of data for "similar" events is considered. It is a known observation that human failure events differ so much in the circumstances and mechanisms leading to failure that each one seems unique, and therefore confining events into a statistical population can produce tolerance problems which are often unacceptable.

Despite these problems, HRA data has been collected since 1960s. Earliest interest in the man/machine interface was mostly confined to the military and the first data sets reflected this perspective. As the nuclear power industry sought to introduce quantitative estimates of human performance into their safety analyses in the 1970s, expert judgement was used to extrapolate the existing body of military experience to commercial nuclear plant operation and resulted in the publication of a THERP handbook /Sw83/ which codifies this extrapolation /Do88/.

The first human reliability methodologies were dominated by the mechanistic view of human performance. In such a behaviouristic view, only the consequences of human errors are accounted for without the reasons for and underlying mechanisms of these errors. Further investigation concerning human performance in the commercial nuclear power plants, especially after the TMI incident in March 1979, caused a change in the perception of the data needs. The model underlying this changed perception considered the human more as a system analyser and decision maker. Simulator training sessions began to illuminate more clearly the role the operator played as an active participant in accident sequence evolution and resolution. These sessions also indicated the potential for gathering high-level decision making data from the simulator.

There were attempts to include some achievements of the psychological theory of cognition in new human reliability analyses /Ra83, Re87/. This implies the necessity of treating human reliability data bases differently and the development of new types of human reliability models

in which phenomenological considerations and cognitive theories are brought into play. Some research programmes have been aimed at verifying human cognitive reliability models based on data from simulators.

The collection of data sets has been greatly increased over the last several years so that now patterns in the data are just being realized /Do88/. An example of such a human reliability related data base is NUCLARR /Ge89/ and the technique developed to extract data of interest and to estimate probabilities of human error intentions applied in the INTENT method /Ge92/. Several research programmes aimed at the collection of human reliability data and verifying human reliability models have been carried out in the USA and other countries /Do88/.

Thus, various data sources, human reliability models and HRA techniques are now available. The problem is how to apply them in the HRA process for a power plant of given design with specific man-machine interface to cover a wide spectrum of possible human interactions at different cognitive levels with underlying error mechanisms and to find out more important human-induced contributors to the risk. It is of paramount importance for decision making in the design and operation of the plant to enable the introduction of appropriate changes in design of the man-machine interface, modified training and improved administrative or operational procedures.

It is generally preferred to use HRA models covering various specific cases with limited contributions of expert judgement. However cases are encountered where extensive expert judgement is required. Several human reliability quantification techniques, more frequently used in HRA/PSA practice, are described below grouped as: time-dependent methods, decomposition methods, expert estimation methods and other methods. They were selected and assessed as candidates to develop relevant software supporting HRA.

2.3.2 Time-dependent methods

2.3.2.1 Time Reliability Correlation (TRC)

A time reliability correlation is one of several psychometric techniques /Do88/. The basic premise of a TRC approach lies in the nature of an offnormal event and the requirement it places upon the operators who must respond to the event. The occurrence of an offnormal event forces the operators to diagnose the situation at hand, interpret its implications on the plant safety, and to decide on a plan to respond - all in the real time of the event phenomena. This event-driven regime of behaviour sets in motion cognitive, perceptual and executive processes that take time to implement. Variability in specific event characteristics, and individual personal characteristics places a fundamental uncertainty on the prediction of a response. It is this source of uncertainty that is observed in simulators as distributed response times and must be quantified with some time-dependent distribution /Do88/.

The fundamental idea of this approach has been adapted to several HRA techniques: the THERP Nominal Diagnosis Model, the Accident Initiator Progression Analysis (AIPA), the Operator Action Tree (OAT) method, the Human Cognitive Reliability (HCR) model, and has been a subject of research in several programmes /Do88/.

Except for the HCR technique, applications of the TRC method are often based on speculation rather than data. The interpretation of data from simulator exercises by LaSalle and ORNL

suggest that the lognormal family of distributions can be also considered in modelling TRCs. A TRC method based on multivariate lognormal distribution has been used in several PRAs to quantify time-dependent human failures. Its formal representation is a random variable T that accounts for the time needed to successfully respond to the situation, i.e., without an unrecoverable mistake /Do88/:

$$T = \tau_R \times \tau_U$$

where: τ_R is a lognormal random variable with median of m and error factor f_R to account for the uncertainty of the process; and τ_U is a lognormal random variable with median of 1 and error factor, f_U to account for the uncertainty in the model.

The complementary cumulative distribution function (CCDF) of this random variable presents the TRC. The formulation as above can be used to obtain a CCDF representing the probability that a response has not been successful by a given time, t i.e. $\text{CCDF} = \text{pr}(T > t)$. An algorithm to compute percentiles of probabilities based on underlying distribution has been incorporated into the ORCA code. Some functions of this code are described in Section 2.3.6.

The data from the simulator exercises seem to indicate that one TRC is not sufficient to accommodate all event and actions types /Do88/. It was proposed to generate a lognormal family of TRCs by a method to "adjust" the TRC parameters based on differing conditions. In adjusting a TRC the slip probability can be used as a threshold value that stops the overall event probability from decreasing below some acceptable lower limit.

The TRC approach is used for quantifying mistakes. Different curves of TRC for the major failure categories: response, response with hesitancy, recovery and recovery with hesitancy are proposed in /Do88/ and presented in Fig. 2.2.

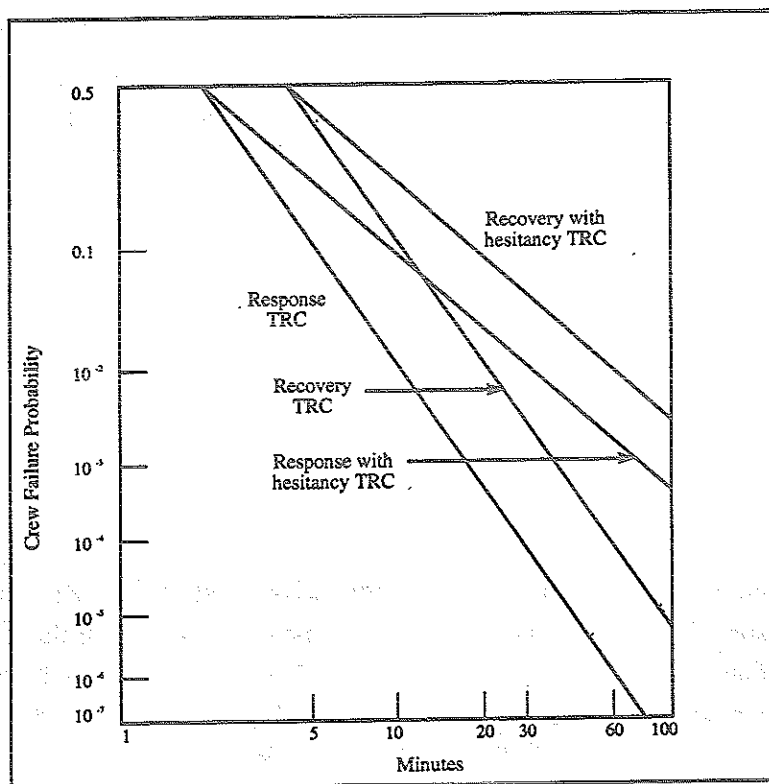


Fig. 2.2: The basic time reliability correlation /Do88/

Remarks:

The TRC concept still needs basic research. More data is needed to provide realistic estimates of the uncertainties associated with TRCs and to investigate the influence of PSFs on probabilistic estimates, as the time-dependence assumption may simplify the representation of operator action to the point at which other PSFs are overlooked. The TRC used in the Operator Action Tree (Section 2.3.2.3) loses most of what may be essential information for some purposes, e.g. it does not make use of much of the information obtained in a full-scale task analysis.

Important issues for further investigation should include the effects of decision making within a crew on TRCs, especially in situations with tradeoffs between competing resources. A method of introducing other influences into a TRC system is needed. One possible approach is the use of a success likelihood index (SLI) to interpolate between CCDF confidence bounds, say the 5% lower bound and the 95% upper bound.

2.3.2.2 Human Cognitive Reliability (HCR) model

The Human Cognitive Reliability (HCR) model /Ha85b/ has been developed for the evaluation of the human non-response probability under time-constrained emergency situations. The quantification of this probability is based on key input parameters, which identify: the type of cognitive behaviour of the operator, the median response time, and some selected PSFs. The median response time ($t_{0.5}$) represents an estimation related to an operator/crew success probability of 0.5, i.e. $\text{pr}(T \leq t_{0.5}) = 0.5$. It is usually evaluated from simulator experiments, task analysis or expert judgement. The performance of the operator is schematized according to the three cognitive processes, identified as skill-, rule-, or knowledge-based behaviour /Ra83/.

The effect of operationally induced stress, experience of operators and quality of operator-machine interface on crew performance is generally expressed by means of a set of PSFs which induce a modification of the median time to perform the related task. The results of the analysis of data for the tasks to be performed have led to the conclusion that the shapes of normalized curves of non-response probability are as illustrated in Fig. 2.3. The shape of the curves was derived from previously available simulator data /Wr82/ and small scale tests /Hu88/. These shapes are reported as approximating a three-parameter Weibull function of the form

$$P(t) = \exp \left[- \frac{(t / t_{0.5}) - C_{e,i}}{C_{g,i}} \right]^{B_i}$$

where: $P(t)$ is the crew non-response probability for a given system time window t , i.e. the time allowed by the system for crews to complete actions before a change in the plant state; $t_{0.5}$ is the estimated median time taken by the crew to complete an action(s) or task(s); $C_{g,i}$, $C_{e,i}$, B_i are the coefficients associated with the predominantly i -th type of mental processing, e.g. skill, rule or knowledge which can be calibrated by simulation data.

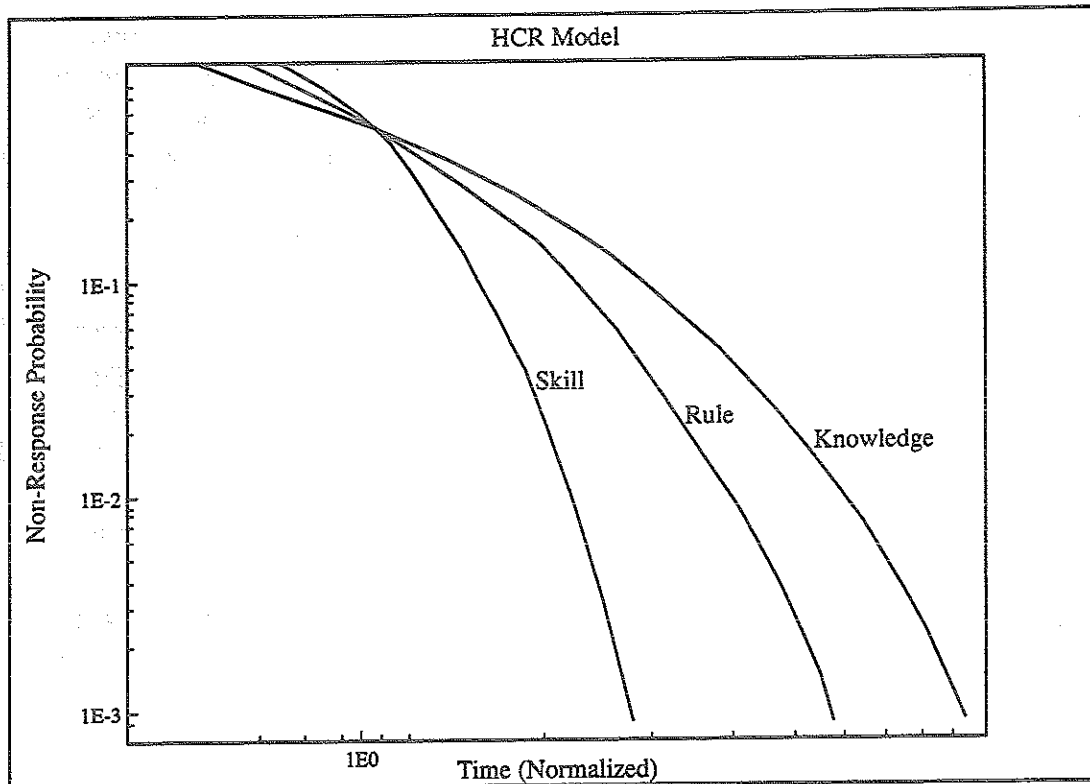


Fig. 2.3: Normalized crew non-response curves for skill-, rule-, and knowledge-based behaviour /Hu88/

The probability of non-response, defined by this equation, becomes an element in the assessment of the overall non-success probability evaluated as input to PSA. The non-response probability is incorporated into a logic representation e.g. human reliability analysis event tree (HRAET) or operator action trees (OAT). Other aspects of human reliability (e.g. selection of a non-viable option) must be assessed by another method to determine the total human error probability.

Remarks:

HRC is a very handy and easily applicable formulation and it does not need any specific mathematical elaboration before evaluating non-success probability. It attempts to cover knowledge-based as well as skill- and rule-based behaviour. The HCR model can be applied in a number of ways. It is possible to apply the hand calculation or a computer program. A logic tree can be provided to assist the analyst in classifying the type of cognitive processing for required tasks /Hu88/.

HCR, being a method based on expert judgement, suffers from the same disadvantages of similar approaches that require extensive and laborious activity by experts. In addition, some recent research on the mechanisms of human performance do not so easily discern between rule- and knowledge-based behaviour, particularly in the light of the new emergency procedures with their symptom-oriented diagnostic "rules" /Do90/.

In practice it is extremely difficult to ascertain whether an action is executed as an immediate response to a particular pattern of indications (skill-based), or it requires the utilizing of an

implicit or explicit rule (rule-based). If an event is assigned to the wrong category, this can have serious consequences, particularly at longer time intervals, i.e. the expected non-response probability can differ by even two orders of magnitude for two categories (Fig. 2.3). Studies of mental processes during actual incidents show that operating teams are constantly changing their levels and modes of processing as an event develops /Re87, Wo87/. Their cognitive behaviour depends upon the level of training and professional experience.

The HCR method assumes that the same form of relationship exists between normalized time and non-response probability for both simulators and real incidents. This is fundamentally untestable because it would not be possible to observe response times for a representative sample of such incidents. The HCR correlation has been shown to be extremely sensitive to estimates of the time available. Thus, the median response time must be estimated accurately. However, it can only be done in practice for a representative, defined accident for assumed initial state and scenario. On other hand, the model is relatively insensitive to changes in the PSFs considered (only a few PSFs can be taken into account).

The development of this model has been funded by the Electric Power Research Institute, which with the NRC and Sandia National Laboratory has jointly used the LaSalle simulator to try to "verify and confirm" the model. No detailed results are known from the EPRI effort, but the Sandia analyses did not indicate any need to use a Weibull distribution family nor any justification for the three types of behaviour to produce characteristic curves /Do88/.

A serious deficiency of the HRA technique is that it does not provide a complete HEP, and it addresses only a limited aspect of operator performance in emergency responses, i.e. the probability that he will not process information related to diagnosis and required action within the time available. A total human error probability can be obtained using e.g. the expanded operator action tree (OAT) representation /Hu88/ or an extended THERP method including TRC of post-diagnosis actions /Re92a/.

2.3.2.3 Operator Action Tree (OAT)

The Operator Action Tree (OAT), proposed by Wreathall /Wr82/, is a method for the representation of the sequence of actions to accomplish a function or task. This approach denotes the fact that purposeful human activity is first stimulated by some event. It can be an urgent message such as an alarm in the control room or a more passive stimulus such as recognition that it is time for periodic test or maintenance of some equipment. Selection of the representation of human interaction begins with the definition of the stimulus, which may be an initiating event for an accident sequence, a manoeuvre for normal operation, planned maintenance, or the diagnosis of a plant state which requires plant shutdown.

OAT focuses on decision making and comprises alternative actions on the basis of ambiguities or operator interpretations associated with observations, diagnosis and selections of required responses. It allows the analyst to identify and graphically represent various types of decision strategies that might affect the accident sequence. The method thus allows for analysing misinterpretation at various stages in the execution of a task.

OAT contains a time failure, or non-response probability relationship, used to quantify operator errors. Various TRC curves are proposed for the quantification of some branches of

OAT, e.g. curves are considered in the SHARP framework /Ha84/ which show the essential character of skill-, rule- or knowledge-based operator behaviour.

In this way it is possible to model various decision paths that the operator may take and the assessor can determine whether they are key decisions or not. The other important aspect is concerned with the evaluation of the limited time which the operator has to carry out the task. OAT contain, therefore, probabilities evaluated using suitable techniques, e.g. TRC technique (OAT-TRC) and other to account for non-response events of interest.

The expanded version of OAT (EOAT) also includes the analysis of the detection stage and thus the possibility of performing a successful or a faulty detection, as well as no detection at all. An example of EOAT is presented in Fig. 2.4.

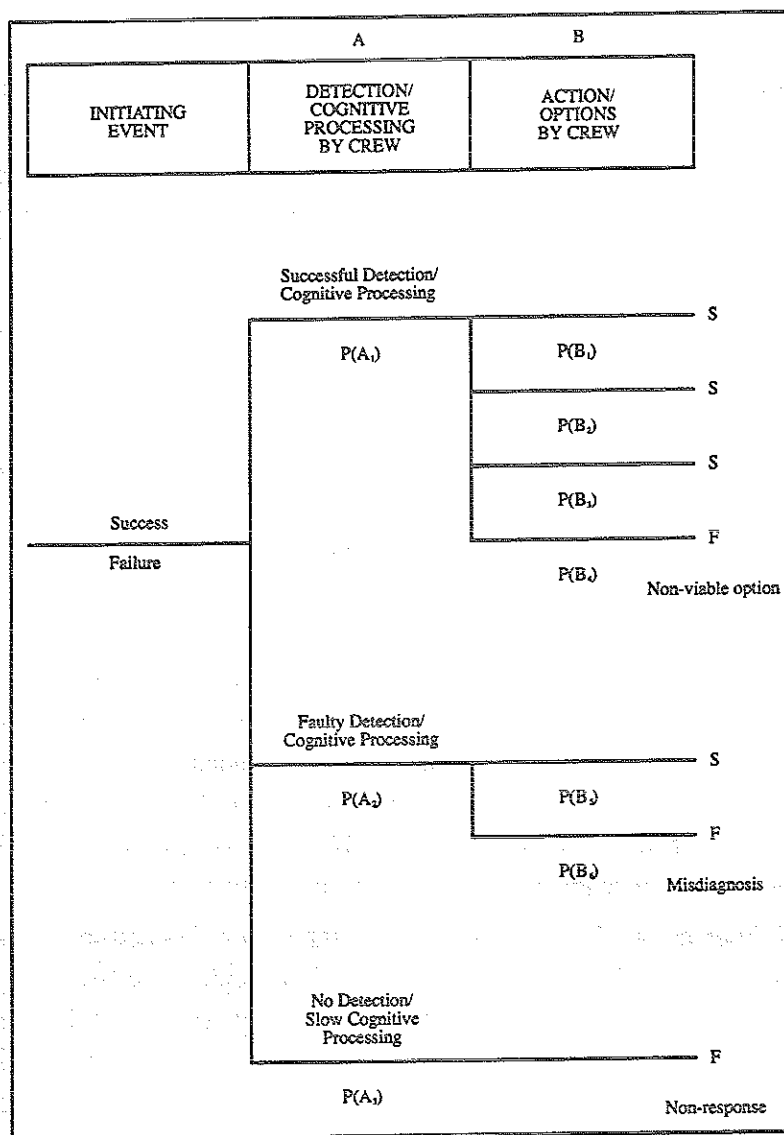


Fig. 2.4: Expanded Operator Action Tree (EOAT) representation /Hu88/

Remarks:

OAT allows the analyst to identify and display various types of decision strategies that may affect the accident sequence conceptually. It is possible to model various decision paths that the operator may take and the assessor can determine whether they are key decisions or not. The other important aspect is concerned with the evaluation of the limited time in which the operator has to carry out the action. The OAT representation may be appropriate for including the recovery actions. Thus, OAT can be useful to analyse some reliability aspects of operator activity, especially in time constraints.

2.3.3 Decomposition methods

2.3.3.1 Technique for Human Error Rate Prediction (THERP)

The Technique for Human Error Rate Prediction (THERP) was developed for the US Nuclear Regulatory Commission /Sw83/ and is the most widely used technique to date. It is basically a hybrid approach because it models human errors using probability trees and models of dependence, but also considers Performance Shaping Factors (PSFs) affecting the operator actions. The technique is dependent on a data base of human error probabilities summarized in Chapter 20 of the THERP handbook /Sw83/. These data were derived from a mixture of objective field data and judgements by the authors of the technique. The data base makes it unique amongst HRA techniques. This, together with its engineering approach, and the fact that it was the first major methodology to be accepted and used in the field, almost certainly accounts for its popularity.

The analyses are carried out in four phases: (1) familiarization; (2) qualitative assessment; (3) quantitative assessment (assigning nominal HEPs, estimating the relative effects of PSFs, assessing dependence, determining success and failure probabilities, determining the effects of recovery factors) and (4) incorporation (performing a sensitivity analysis, supplying information to system analysts).

Once the system failure of interest, i.e. failure of the hardware that may induce and influence human errors, has been identified, a detailed task analysis and human error analysis is performed. The implicit classification of human errors, included in the data base of THERP, contains errors that are all reducible to two error types: (a) Error of Omission (EOM), by which a step or an entire task are omitted; and (b) Error of Commission (ECOM), which entails selection error, sequence error, time error, or qualitative error /Sw83/.

The failure probabilities are evaluated by combining the data contained in the data bank of the Handbook (Chapter 20) and data derived from recorded incidents, trials from simulators and subjective judgements. However, the major source of information is contained in the data bank itself, because it consists of data tables of human error probabilities and associated factors as well as performance models explaining how to modify the data for changes in performance shaping factors and guidelines on how to convert independent failure probabilities into conditional failure probabilities. An expert system for the application of THERP technique to support PSA is described in item 3.5.2.

In order to describe and analyse these human operations and errors, THERP uses the Human Reliability Analysis Event Tree (HRAET) as its basic tool (Fig. 2.5). A graphical description of

the procedural step in a task is set out in a logical framework by HRAET. At each node of the tree there is a binary decision point, representing the failure or the success of the current action. Since these trees are compatible with conventional system event trees, they can be evaluated in the formal mathematical sense, and consequently, if the success or failure probability of each particular task step in a procedure is known, the overall reliability of the procedure can be calculated (Fig. 2.5). This decompositional approach is very similar to the engineering risk assessment of hardware components.

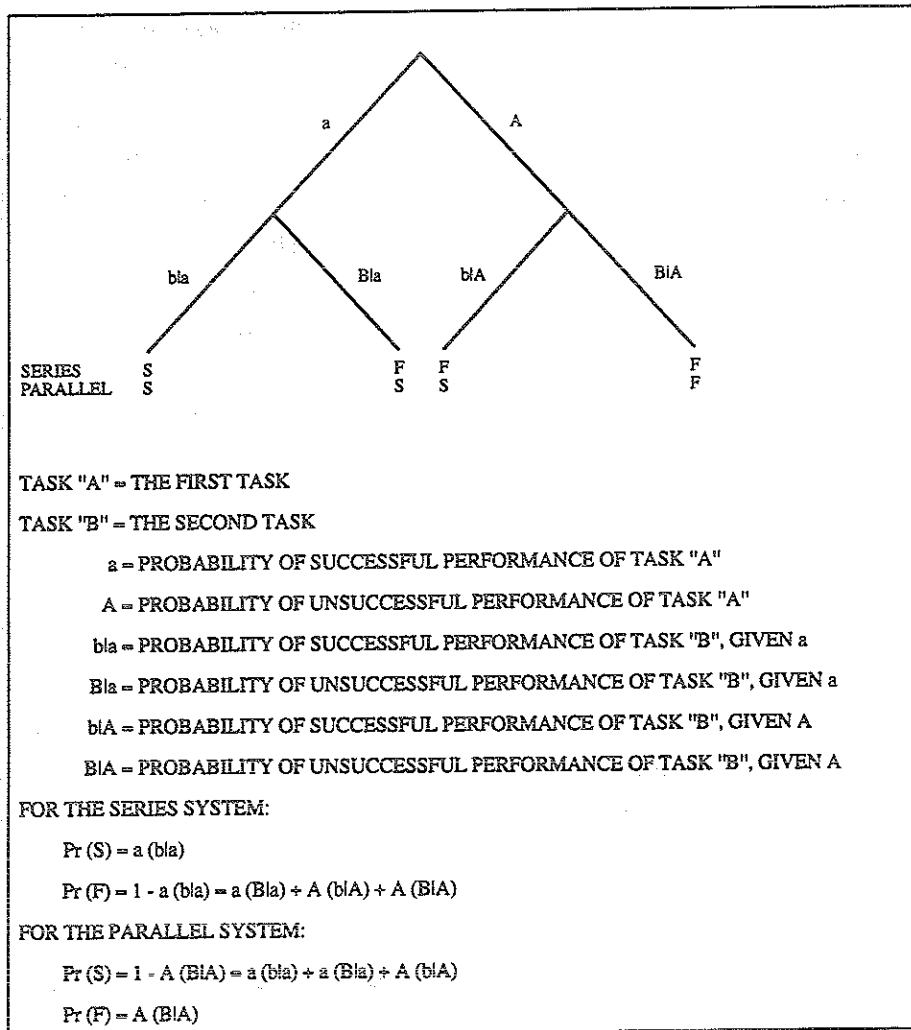


Fig. 2.5: Human Reliability Analysis Event Tree (HRAET) for series and parallel systems

In the analyses it is important to account for the dependency of various tasks in accordance with the guideline outlined. Five levels of dependence are possible, from "zero dependence" to "complete dependence". Opportunities for recovery of human error should also be identified. The consideration of error recovery is important as it may significantly reduce the overall human error probability for a task. The probabilistic evaluation is synthesized using conventional probabilistic mathematics for the combination of HEP values obtained for each task step. The overall task error probability is then incorporated in the system reliability assessment.

Remarks:

THERP provides a structured, logical and well represented record of factors and errors considered in human reliability assessment, which makes it very useful for a reference basis and as a data source as well as for the sensitivity analysis. Moreover, the compatibility of the methodology with fault/event tree methods allows THERP to be integrated into PSA. The level of detail of the THERP analysis can be tailored according to the depth of the overall assessment being performed, e.g. small effort initial assessments would only require a crude THERP analysis. THERP can be used either with the handbook data or any other relevant data, e.g. simulator data, known plant-specific data or data obtained from other HRA methods.

On the other hand, this technique does not allow the evaluation of errors concerning post-diagnosis or high level decision making. Moreover, THERP turns out to be a quite rigid tool, i.e. it is not particularly sensitive to many possible PSFs nor does it attempt to identify fully the underlying psychological aspects of a particular error. The human reliability data and probabilistic distributions proposed in the handbook are not systematically validated.

A detailed THERP analysis requires a considerable effort to obtain HEP values. It seems, however, that deep decomposition of human actions into very detailed tasks/errors is not systematically validated, either. Peer review problems have been encountered during the peer review studies, particularly related to a lack of uniformity in the modelling of the scenarios and the use of the data base. This led in some cases to large discrepancies in estimates evaluated by different analysts for the same task.

2.3.3.2 Accident Sequence Evaluation Procedure, Human Reliability Analysis Procedure (ASEP-HRAP)

The Accident Sequence Evaluation Procedure, Human Reliability Analysis Procedure (ASEP-HRAP) /Sw87/ is a shortened version of the THERP handbook /Sw83/. The intention of this procedure was to enable systems analysts to make estimates of human error probabilities with minimal support from experts in human reliability analysis, to make estimates of human error probabilities and other human performance characteristics which are sufficiently accurate for probabilistic risk assessments /Sw88/.

The basic approach taken in the development of the ASEP-HRAP was to select certain generic HEPs for certain sets of tasks and to employ easy-to-understand procedures for using these HEPs and for estimating the effects of dependence and recovery factors. The goal was a rule-based procedure that could be employed with considerably less judgement than is the case with the more complete HRA procedure. To achieve this goal, and to avoid undue optimism in estimating the effects of human error, several conservatisms were employed /Sw87/.

The ASEP-HRAP consists of a Pre-Accident Screening HRA, a Pre-Accident Nominal HRA, a Post-Accident Screening HRA and a Post-Accident Nominal HRA. The screening analysis uses conservative, i.e. pessimistic estimates of HEPs, response time, dependence levels, and other human performance characteristics.

Remarks:

This method can be used by systems analysts with some guidance from specialists in HRA, especially in the difficult area of diagnosis errors, to ensure that undue optimism is avoided. The ASEP-HRAP should be regarded as a first-cut or approximate HRA, which will generally provide conservative estimates of the human contribution to pre-accident and post-accident unavailabilities or failures. The THERP/Handbook approach, when applied by a fulltime team of systems analysts and a human reliability analyst, will then permit a more detailed and accurate systems assessment of the impact of potential human errors. It can be useful to apply the more detailed HRA procedure to those human/system interactions that are assessed as more significant to the system analysed, based on the HEPs obtained using the ASEP-HRAP.

2.3.3.3 Human Error Assessment and Reduction Technique (HEART)

The Human Error Assessment and Reduction Technique (HEART) has been developed by Williams /Wi85, Wi88/. Essentially it considers particular ergonomic and environmental factors (called EPCs - error-producing conditions) which can negatively affect performance. The extent to which each factor independently affects performance is quantified, and the human error probability is calculated as a function of the product of those factors identified for a particular task.

The HEART table of generic tasks consists of nine tasks denoted from A to M. For each task the nominal human unreliability is proposed with 5th and 95th percentile bounds. For the task analysed the generic task is selected with given probability value. Then the error-producing conditions are evaluated. For each factor considered the effect is quantitatively assessed using a simple formula taking into account a HEART effect and a proportion of effect. The assessed nominal likelihood of failure is calculated as a product of the nominal human unreliability and the assessed effect for each factor considered. If the result obtained exceeds 1.00, the probability of failure has to be assumed to be 1.00.

The special table of error producing conditions and associated multipliers consists of 38 positions which are divided into two groups: 1-17 with large effects on reliability (the HEART predicted effect between 17 and 3) and 18-38 with low effects on reliability (the HEART predicted effect between 2.5 and 1.02).

Remarks:

The technique is quick and simple to use. It is also useful in the design process enabling an engineer/designer to link ergonomics directly to risk assessment and/or design, in a quantitative and hence auditable manner. It also enables the engineer/designer to make cost-benefit evaluations, which is a highly useful facility.

However, at present stage of the development, HEART can only be used to assess tasks in isolation, i.e. no model exists of dependency between performing (or failing to perform) a succession of tasks. This could present problems, e.g. in an event tree analysis of operator error. As the results are significantly influenced by judgement and data it would be beneficial if the source data could be reviewed by ergonomist peers. The important issue concerns the independence of the effects of factors influencing human unreliability. This means that there is

no additional effect on performance other than that assigned to a factor by its multiplier, no matter what combination of factors are present. Thus, the technique requires further development.

2.3.3.4 INTENT

INTENT is a new method for estimating human error probabilities for decision-based errors /Ge92/. This method is aimed at identifying and probabilistically quantifying errors of intention. Errors of intention (EOI), an important subset of errors of commission, are related to human cognitive functions. However, an omission could be caused by an EOI, too. Errors of intention may be manifested immediately or can lie dormant until triggered by some insidious combination of hardware and human actions. The cognitive nature of errors of intention and the fact that they result from a wide range of factors, such as poor training or poor safety culture at a facility, make it difficult to model and quantify them.

The identification of the errors of intention relied on applying SNEAK and HSYS to two sources of information, the Nuclear Computerized Library for Assessing Reactor Reliability (NUCLARR) and the Licensee Event Reports (LERs) /Ge92/. The search was conducted to determine the extent to which existing data for errors of intention might already be available in the open literature. In the approach /Ge92/ a list of 20 decision-based errors was generated and an expert judgement technique is outlined for assessing the likelihood of these errors occurring under various conditions. These conditions were defined by different levels of PSFs.

Four source categories for errors of intention were identified: (1) action consequence, (2) attitude leading to circumvention, (3) crew response set, and (4) resource dependences. For each error of intention the upper bound of HEP and the lower bound of HEP were proposed. The lower bound refers to situations where PSFs have been estimated to be optimal. The situation-specific HEP lies between these bounds.

Quantification of the errors of intention, identified in the NUCLARR and LER data sources, were performed applying the SLIM technique based upon group consensus of both relative PSF importance and PSF adequacy for given scenario. Eleven PSFs were taken into account: human-machine interface; skill-, rule- and knowledge-based behaviour; stress; experience; safety culture; training; motivation; workload; supervision; communication; and procedures. A calculation method is proposed to assess the plant-specific HEPs.

Remarks:

The method seems to provide data which can serve to remedy a major deficiency present in PRAs conducted to date, i.e. to account for rare, high consequence failures due to errors of intention committed during various plant evolutions. The generated list of errors may not be exhaustive, but, it does provide a foundation on which to build a more complete data base as field data becomes available.

The authors of this new method suggest that future research should attempt to: (1) extend the variety of errors of intention identified to date, (2) determine whether the suggested weights of PSFs will hold true for these 'new' errors of intention, (3) apply the INTENT method in an

HRA program, and (4) attempt to tie the data and categories to some high level cognitive theory (e.g. GEMS /Re87/).

2.3.4 Expert estimation methods

2.3.4.1 Success Likelihood Index Method (SLIM)

The Success Likelihood Index Method (SLIM) originates from the field of decision analysis, and it is basically a method for quantifying the preference in a set of options /Em83, Em84/. Its applicability to the assessment of human reliability derives from the consideration that human performance is affected by many factors and thus it can only be quantified by calculating the additive effects of these factors on human response. Such PSFs as the time available to perform a task, quality of procedure, level of training etc. can be considered. A systematic approach is used to identify these PSFs for a specific set of tasks. The relative importance (weights) of the PSFs are derived by structured expert judgements.

The method comprises two modules: SLIM-MAUD (Multi-Attribute Utility Decomposition), for quantifying the effects of various factors on human reliability, using expert judgements for deriving the relative likelihood of success for a set of tasks; and SLIM-SARAH (Systematic Approach to the Reliability Assessment of Humans), for transforming these relative likelihood into absolute probabilities, using log-linear calibration relationships.

The SLIM procedure comprises 10 steps: (1) Definition of situations and subsets; (2) Elicitation of PSFs; (3) Rating of the tasks on the PSFs; (4) Ideal point elicitation and scaling calculations; (5) Independence checks; (6) Weighting procedure; (7) Calculation of the Success Likelihood Index; (8) Conversion of the SLIs to probabilities; (9) Uncertainty bound analysis; (10) The use of SLIM-SARAH for cost-effectiveness analyses. The first eight steps are performed in the SLIM-MAUD module and the remaining steps are carried over by SLIM-SARAH.

The first phase of applying SLIM-MAUD involves defining the situations to be evaluated by providing the judges with as much information as possible regarding the characteristics of the tasks, the individuals who will perform them, and factors that will influence the likelihood of success. The tasks are divided into subsets for which the task success probabilities are influenced by common sets of PSFs. Each subset is then assessed in a separate SLIM-MAUD session.

After that the judges are asked, e.g. during the interactive session at the computer, to identify and to rate the PSFs relative to the tasks, on the basis of a linear scale from 1 to 9. The judges have also to define the ideal point on the particular scale being used, and this allows MAUD to scale all the other ratings. The rescaled ratings for each task on the various PSF scales are subsequently multiplied by their respective relative importance weights and this products are summed to give the overall SLI for each task.

The weighting procedure is concerned with evaluating how much emphasis is to be given to each of the PSF in terms of its effect on the likelihood of success. The SLIs for each task are calculated using the following formula:

$$SLI_j = \sum_i W_i R_{ij}$$

where SLI_j represents the success likelihood index for task j ; W_i is the normalized importance weight in the i -th PSF (the weights for all the PSF's sum to 1); and R_{ij} is the scaled rating of task j on the i -th PSF (the rating is scaled in terms of its distance from the ideal point).

The SLIs generated in a SLIM-MAUD session are relative measures of the likelihood of success of each of the tasks considered in the session. In order to transform these to human error probabilities it is necessary to calibrate the SLI scale for each set of tasks considered. The relationship assumed in the SLIM method is:

$$\log HEP = a SLI + b$$

where a and b are constants. Determination of the constants in this 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. This produces two equations from which the constants a and b can be calculated. Several possible sources for calibration data exist, e.g. data from training simulator sessions or absolute probability judgements /Hu88/. The justifications for this formula are partly related to the Paired Comparison (PC) technique, which also makes use of experts' judgements of likelihood of success, as well as more theoretical work supporting the validity of the logarithmic relationship.

Remarks:

The method for the calculation of SLI is based on a good theoretical background in decision theory and the easily managed interactivity with the computer makes the technique highly scrutable and auditable. Once the detailed data base has been established, the evaluation can be made easily and rapidly. The SARAH module allows relatively powerful cost-benefit design questions to be answered quantitatively, so that the impact of a design change on the human reliability for a task can be quickly calculated. SLIM-MAUD does not require extensive decomposition of a task to an elemental level (as THERP does), but is capable of quantifying human reliability at a higher or more holistic level of task description.

The major drawback of the SLIM method is that it makes extensive use of expert judgement requiring a small group of (e.g. 4) experts. The technique requires rather the use of a microcomputer to support the calculations. As with the paired comparisons technique, the logarithmic calibration equation would benefit from more extensive empirical and experimental support. For assessment of uncertainties the Absolute Probability Judgement (APJ) method is proposed /Hu88/.

2.3.4.2 Socio-Technical Approach to Assessing Human Reliability (STahr)

The Socio-Technical Approach to Assessing Human Reliability (STahr) technique /Ph89/ is strongly influenced by the SLIM-MAUD approach. The STahr consists of a "technical component", an influence diagram, and a "social component". The technical component involves developing a description of the "target event" (task) as well as the general setting and conditions leading up to the event. The modelling technique of the influence diagram, borrowed from decision theory, is potentially easier to use than event trees or fault trees. Modelling only dependences between events, the influence diagram organizes them as a system of conditional probabilities. An example of the influence diagram is shown in Fig. 2.6.

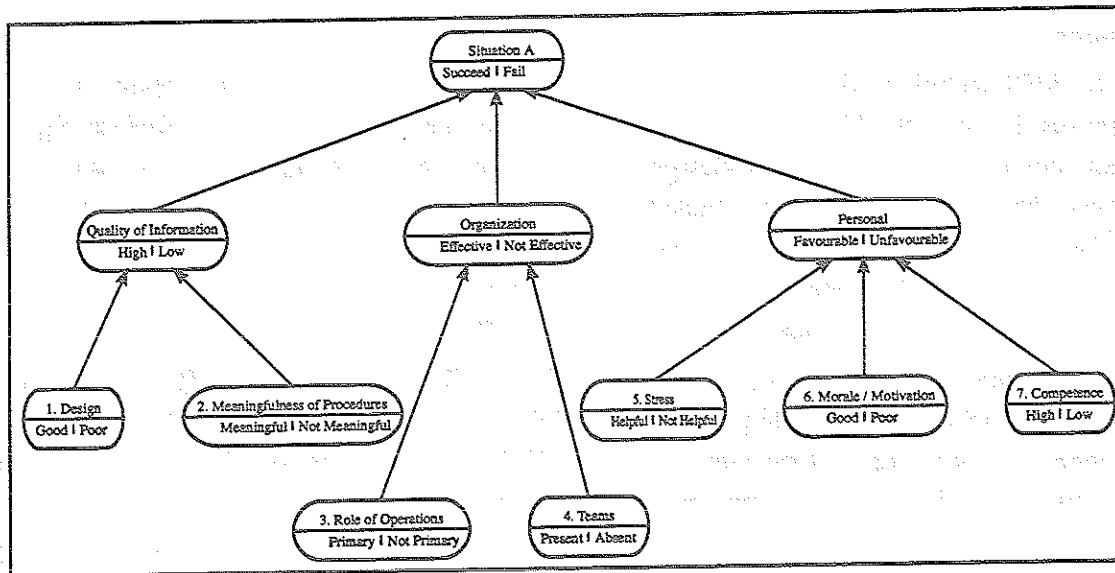


Fig. 2.6: The STAHR influence diagram /Ph89/

The remainder of the method relies on the knowledge and judgement of several subject-matter experts. Experts from several domains should be represented depending on the analyses and assessments of interest and organizing procedure, e.g. group consultant and facilitator, technical moderator, reactor operators trainer, thermohydraulic engineer and procedures specialist, probabilistic risk analyst, reliability and systems analyst, human reliability analyst and reactor operator.

Using the influence diagram is a matter of applying the following ten steps (but not necessarily in this order): (1) Describe all relevant conditioning events; (2) Define the target event; (3) Choose a middle-level event and assess the weight of evidence for each of the bottom influences leading into this middle event; (4) Assess the weight of evidence for this middle-level influence conditional on the bottom-level influences; (5) Repeat steps 3 and 4 for the remaining middle and bottom influences; (6) Assess probabilities of the target event conditional on the middle-level influences; (7) Calculate the unconditional probability of the target event and the unconditional weight of evidence of the middle influences; (8) Compare these results to the holistic judgements of the assessors. Revise the assessments as necessary to reduce the discrepancy between holistic judgements and model results; (9) Iterate through the above steps as necessary until the assessors have finished refining their judgements; (10) Do sensitivity analyses on any remaining group disagreements. Report either point estimates if disagreements are of no consequence, or ranges if disagreements are substantial.

The group of experts participates in assessing the "weights of evidence" for various levels of influences on adjacent levels and on the target event. The target event is then assessed, conditional on the middle-level influences. The unconditional target event probability and the unconditional weight of evidence of the middle influences are then calculated with a computer. These results are compared with holistic judgements made by assessors. Any discrepancies are discussed and revised. The process outlined above is then repeated as necessary until the assessors have finished refining their judgements.

Remarks:

The STAHR technique depends on the coalescing of many experts' diverse opinions into a consensus judgement. Although this process at times can be difficult and frustrating, the authors contend that "diversity of viewpoint" is a key criterion in composing the groups. They estimate that at least one day is needed to train participants in the STAHR approach, and building confidence and familiarity takes another day. Then, new target events take no more than two hours for the complete set of the assessments, and variations on already-completed target events take less than one hour to complete /Ph89/.

There are some controversial trials comparing the STAHR approach with THERP. Although STAHR requires several participants, who usually go back to first principles, THERP is considered to give only approximate results. There is also the opinion that successful application of THERP also requires several experts to be involved /Ph89/.

The STAHR approach has potential for assessing human reliability in complex systems. By concentrating on influences rather than individual events, it avoids the bottom-up detail of event trees and fault trees with their attendant bushiness in attempting to cover all event-sequences that could possibly happen, and the high costs of applying those approaches for complex systems.

Unlike handbook-based approaches, STAHR cannot be applied simply by following some established steps in a procedure. Rather, considerable flexibility is built into STAHR, and participants are forced to think about fundamental aspects of the operation of the system including human aspects. Because of this, probabilities for target events cannot be developed quickly and at a low cost.

The STAHR approach is still at the early development stage, so it is not possible to judge its utility fairly. However, as with any expert judgement method, the inconvenience of using a large, diverse group of subject-matter experts will often eliminate the consideration of using this technique for human reliability evaluation.

2.3.4.3 Pair Comparison (PC) technique

The Pair Comparison (PC) technique /Se83/ is a structured expert judgement method for quantifying human error probabilities. It does not require any quantitative assessment. The experts are rather asked to compare a set of pair of tasks for which HEPs are required. For each pair, the expert must decide which has the highest likelihood of error. The technique elicits these comparative judgements from a number of experts and develops a scaling of the tasks in terms of their relative likelihood of error. Two or more tasks, with known HEPs, are then used to calibrate this scaling, based on a logarithmic transformation to derive HEP estimates.

The complete PC procedure for estimation HEPs is developed in 16 steps. The procedure and probabilistic evaluation is quite complex. The method requires a considerable number of experts to be involved. It is recommended that the number of experts should preferably be more than ten. Although it is possible to use a smaller number of experts, the calculations become more difficult, and the derived scale will be less robust /Hu88/.

The Paired Comparison technique is considered by some researchers as an approach which does not require the experts to be involved to have a great deal of expertise in probability or mathematical and statistical concepts. Such an opinion can be questioned when the PC methodology is examined from the point of view its application for human reliability evaluation in the case of multidimensional events involving complex human activities with some dependences in comparisons made.

Remarks:

The PC technique enables estimates of probabilities to be obtained using statistical methods applied to comparative estimates with regard to calibration points. In this method a minimum of two empirically estimated HEP values are required, thus the technique can be useful even in cases of scarce empirical data. The probabilistic estimate extracted from the comparative judgements seems to be less biased than that obtained from direct numerical assessments, especially if individual judges are poorly qualified. With a small number of tasks, and a set of readily available experts, the technique can be applied fairly quickly if carried out on a computer.

Moreover, even without calibration, the technique is useful for scaling the relative importance of different human errors and this rank ordering can be helpful in quantifying probabilities using other methods. Experts do not have to carry out the comparison as a group, and this eliminates the psychological problems and drawbacks identified in the case of APJ, although they may still need to be trained as a group.

However, the PC method relies on three main assumptions, which can be violated in the human reliability evaluation, namely: the small complexity of the event considered, the homogeneity of the events or tasks to be performed, and the independence of each comparison made. The PC technique was originally developed to consider relatively simple perceptual phenomena, usually considered as "uni-dimensional" events. When the experts are asked to compare two complex operator error scenarios, such as during NPP emergencies, a simple comparison may prove difficult due to the multidimensional nature of two situations. Similarly, when the tasks to be performed are not homogeneous, i.e. when they are not all maintenance tasks or emergency actions, then the comparison process is very difficult with a potential for substantial errors. The independence criterion aims to identify tasks to be compared which do not create biases for the experts, because these tasks can be somehow related to previously examined cases and that would distort the results of comparisons.

Thus, the PC method may not be appropriate for evaluation of more complex human failure events and related HEPs, especially in situations when the influence of several PSFs should be considered.

2.3.4.4 Absolute Probability Judgement (APJ)

The Absolute Probability Judgement (APJ) /Se83/, also known as Direct Numerical Estimation, is the most direct approach to the quantification of human error probabilities

(HEPs). It relies on the utilization of experts to estimate HEPs on the basis of their knowledge and experience. The rationale for using expert judgements in human reliability assessment is that there are little or no relevant human error probability data, while there exist experts who have the experience and appropriate knowledge that can be translated into quantitative estimates of probability of occurrence of an event.

There are two forms of APJ, namely: Group APJ and Single Expert Method. The use of group methods involves more research, since it is seldom the case that a single expert has enough relevant information and expertise to accurately estimate human reliability. In the group methods, the single opinions based on knowledge of each expert 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 human error probability possibly with some reference to a data base. This latter approach is frequently used both in human reliability assessment and in hardware reliability assessment, though it is arguably a less reliable approach /Hu88/.

There are two primary requirements of any expert: (i) She/he must have substantive expertise, i.e. the expert must know the problem sufficiently well; (ii) He/she must have normative expertise, i.e. the expert (alone or with the aid of a facilitator) must be able to accurately translate this expertise into probabilities.

If an expert does not possess substantive knowledge, assessments will be based only on random guessing. With substantive knowledge but no normative expertise, the expert will be unable to accurately translate his/her beliefs and knowledge about human actions into probabilities.

If a single individual expert approach is being utilized, there is an obvious reliance on one person's expertise, which may be biased, or not comprehensive enough to generate an accurate estimate. It is therefore recommended that a group of experts should be used where practicable. There are four major group methods specified in order of their overall preference /Hu88/: (i) Consensus Group Method; (ii) Nominal Group Technique; (iii) Delphi Method; and (iv) Aggregated Individual Method.

The overall APJ procedure foresees the following steps: (1) Select subject matter experts; (2) Prepare task statements; (3) Prepare response booklets; (4) Develop instructions for subjects; (5) Obtain judgements; (6) Calculate inter-judge consistency; (7) Aggregate individual estimates; and (8) Estimate uncertainty bounds.

Some researchers suggest that six experts would be sufficient for direct estimation, although more would be preferable. In practice however, financial and other constraints often lead to the use of a smaller group of only three or four experts. Obtaining evaluations from experts varies for the methods mentioned above:

i	In the Consensus Group Method each member contributes to the discussion, but the group must arrive at an estimate upon which all members of the group agree;
ii	In the Nominal Group Technique, which is similar to Delphi Method described below, except that some limited discussion is allowed between experts for clarification purposes only;
iii	In the Delphi Method experts make individual assessment but then they can review and reassess their individual estimates on the basis of the other estimates which are shown to everyone; the revised estimates are then statistically combined as described for the method below;
iv	In the Aggregated Individual Method experts don't have to meet, but make estimates individually and then these estimates are combined statistically by taking the geometric mean of the individual estimates.

Tab. 2.5: Methods for obtaining evaluations from experts /Hu88/

When non-consensus methods are used, a statistical procedure for calculating the inter-judge (step 6) is applied, based on a technique of the analysis of variance (ANOVA). If the agreement between judges is adequate, the aggregation of the individual estimates is done (step 7) by taking the geometric mean of the single estimates. Finally the uncertainty bound is estimated (step 8) using the statistical concepts of standard error and confidence range.

All forms of APJ, whether group or individual, are prone to particular biases which can occur and detract from the accuracy of the expert assessments. One of the major ways to reduce the problem of biases in expert judgement during an expert group session is to employ the use of a "facilitator", who will generally manage an expert panel session without providing any quantitative estimates.

There are several known biases, the most important among them are: the overconfidence bias, the availability bias and the anchoring bias /Hu88/. Some other aspects of biases in quantitative judgements and ways to avoid them are proposed /Po82/. Judging probability is a special case of quantitative judgement which requires a special practical approach to avoid biases /Ka72/. A heuristic for judging frequency and probability has been proposed /Tv73/.

Remarks:

The APJ technique has been shown to provide accurate estimates in several domains including human reliability assessment /Hu88/. Some studies give support for the validity of this method. The method is relatively quick to use if the data acquiring process is well organized and the experts possess a high level of expertise. This process and discussions between experts, if documented, can often be useful for further qualitative analysis of results obtained.

This technique is prone, however, to certain biases and to personality/group psychological problems and conflicts, which can, if not effectively countered (e.g. by the use of a facilitator), significantly affect the validity of the technique. Since the technique is often linked to 'guessing', it is characterized at a somewhat low level in terms of apparent or 'face' validity.

2.3.4.5 Confusion Matrix (CM) technique

The Confusion Matrix Technique /Po81/ is used to estimate the probability of initiating event misdiagnosis. Actions taken during an event sequence to supplement automatic system response are of major concern. Actions that fall outside the response envelope, such as errors made in test and maintenance activities, are not modelled explicitly in a confusion matrix. The initiating events that are used for this technique must involve accidents that are similar in appearance to the operators.

The creation of a confusion matrix begins by listing initiating events vertically and horizontally. An example of the confusion matrix is presented in Fig. 2.7. The horizontal column represents the actual plant condition and the vertical list represents what the crew might perceive as an event. Thus each matrix element corresponds to a potential misdiagnosis.

Several matrices can be constructed in this manner, for instance, two matrices for two different times in the accident sequence. For example, one might be for 15 min after the first indication, and the second for an hour after. Next, a frequency analysis is carried out for each "actual initiating event" and an estimated total probability for misdiagnosis for each event for both times is calculated.

The confusion matrix approach provides a rationale for representing operator actions which stem from the response to the not correctly diagnosed initiating event/accident. This delineation is a key input for SHARP, Step 5 where the impact of human interactions on the course of the plant accident sequence can be examined to model alternative accident sequences.

Expert opinion is used to rank each initiator in the order in which they would most likely be confused with the actual event by considering the similarity of symptoms and how frequently they might occur. Once the ranking is completed, preferably by a group of experts, probabilities are assigned for each misdiagnosis. The probabilities are then modified to reflect some PSFs, e.g. the quality of the control room design and of operator training.

Probabilities of a confusion matrix can be evaluated with the contribution of experts using one of the multiple-experts techniques, e.g. SLIM, STAHR or APJ, with the relevant procedure. These evaluations can be supported by data from other information sources, e.g. probabilities based on simulator data.

It would be desirable to support assessments with more objective data, e.g. similarity measures evaluated on the basis of curves in time of important plant parameters (symptoms) obtained from the plant simulation for initiating events selected and representative accident scenarios. This stems from the fact that probabilities of misdiagnosis errors will be higher in cases of similar plant responses for different initiating events.

ASSESSED EVENTS POSSIBLE MISDIAGNOSIS	1	2	3	4	5	6	7	8
	SMALL LOCA	LARGE LOCA	STEAM GEN. TUBE RUPTURE	STEAM LINE BREAKS			EXCESSIVE FEEDWATER FLOW	LOW FEEDWATER FLOW
				SMALL INSIDE CONTAINMENT	LARGE INSIDE CONTAINMENT	LARGE OUTSIDE CONTAINMENT		
1 SMALL LOCA			P3,1	P4,1				
2 LARGE LOCA			P3,2	P4,2	P5,2	P6,2	P7,2	
3 STEAM GEN. TUBE RUPTURE	P1,3	P2,3					P7,3	
4 SMALL STM. LINE BRK. INSIDE COM.	P1,4	P2,4	P3,4					
5 LARGE STM. LINE BRK. INSIDE COM.		P2,5		P4,5				
6 LARGE STM. LINE BRK. OUTSIDE COM.				P4,6			P7,6	
7 EXCESSIVE FEEDWATER FLOW	P1,7	P2,7	P3,7	P4,7	P5,7	P6,7		
8 LOW FEED- WATER FLOW				P4,8				
9 LOSS OF OFF- SITE FLOW								
10 NO DIAGNOSIS IS MADE	P1,10	P2,10	P3,10	P4,10	P5,10	P6,10	P7,10	P8,10

P(i,j) = POSSIBLE MISDIAGNOSIS TO BE ANALYSED

i: ASSESSED EVENTS

j: MISDIAGNOSIS EVENTS

Fig. 2.7: Example of confusion matrix (CM) for misdiagnosis of an event /IA91c/

Remarks:

The Confusion Matrix Technique allows only the probability of initiating event misdiagnosis to be estimated. The application of CM is associated with a considerable effort requiring in addition multiple expert methods for the estimation of probabilities of the confusion matrix. It causes the same inconveniences and is characterized by drawbacks of the multiple-expert techniques mentioned above.

A difficult problem is to construct confusion matrices for different times in the accident sequence, associating symptoms with the wrong initiators. This is due to the complexity and psychological nature of the diagnosis process to be performed in time by operators who could attempt to identify the accident situation and to make decisions based not only on its symptoms.

2.3.5 Expert estimation methods seldom applied in PSA studies and other conventional HRA methods

2.3.5.1 MAintenance Personnel Performance Simulation (MAPPS) model

The MAintenance Personnel Performance Simulation (MAPPS) model was developed by Siegel /Si84/. MAPPS is a tool for quantitative analyses of the maintenance tasks. The analyst can vary systematically, individually and in any combination, a variety of conditions, such as task variables, levels of operator ability, environmental and situational conditions, as well as human factor variables. The objective is to yield subtasks, shift, iteration and summary data, which show the effect of the introduced variations on the simulated task performance. The results can be used for defining trade-offs and regulatory decisions.

In order to perform the simulation, the analysts must first perform the task analyses to determine the tasks required and identify the associated subtasks for the individual(s) involved. For each subtask, input data based on ratings (e.g. various PSFs) or actual measurements (e.g. performance times) are entered along with selected parameter values (PSFs). A computer then simulates the performance of each subtask using the model algorithms. There is the possibility of using a Monte Carlo simulation approach to introduce chance elements. The output of the evaluation includes probability of success, time to completion, areas of operator overload, idle time, and level of stress. Using an iterative method, the analyst can forecast the results of a potential design improvement prior to implementation /Sa87/.

Remarks:

MAPPS is thus a model devoted to many users, from power plant architects and engineers to plant maintenance management experts, as well as the regulatory body.

The MAPPS technique represents an attempt to model human performance for maintenance tasks using a set of PSFs that interact with each other, and incorporates recovery factors such as the effects of rest and checks enforced by quality-control polices. The model is computer-based, which allows the analyst to simulate the same task repeatedly and assess how the inputs affect outputs. Numerous inputs, many of which involve subjective ratings, are required from the analyst. Default values are provided for unknown PSFs.

The opacity of the simulation algorithms weakens the face validity and can lead to a lack of confidence on the quantitative values of probability obtained. The basic problem associated with using this model for human reliability evaluation is its validation.

2.3.5.2 TESEO

The TESEO (Tecnica Empirica Stima Errori Operatori) model, developed by Bello and Colombi /Be80/, allows human reliability to be predicted as a function of five factors assumed to be the major determinants of operator performance, namely: (1) The type of activity to be carried out (K_1); (2) The time available to carry out this activity (K_2), also called the temporary stress factor; (3) The characteristics of the operator (K_3); (4) The emotional state of the operator (K_4), called the activity's anxiety factor; and (5) The environmental ergonomics characteristics (K_5), called the activity's ergonomic factor.

The HEP for a given task is then calculated as:

$$HEP = K_1 \cdot K_2 \cdot K_3 \cdot K_4 \cdot K_5$$

The values of each factor can be determined from tables. Whenever the result of the multiplication is greater than 1, it is assumed that $HEP=1$.

Remarks:

The technique is quick and easy to use. It also facilitates sensitivity analyses, i.e. it can be easily used to determine the effect of human factor improvements on human reliability. However, there is a lack of theoretical foundation associated with this technique. The model assumes that the five factors are always adequate for assessment of the human performance and no theoretical justification is given for the multiplicative relationship between these factors and HEP.

2.3.5.3 Other conventional HRA methods

Other conventional HRA methods exist, e.g. AIPA /Fi75/, SHERPA /Em86/ and VISHN /Po88a/. These methods are seldom used in PSA and therefore will not be considered in detail in this report. Other approaches and tools in human reliability analysis are discussed in reference /Ha88/, which includes the CSE (Cognitive Sub-Element) model. The CSE model has been utilized as a part of the operator reliability experiments, together with investigation of the HCR model.

Some approaches are proposed based on known HRA techniques but with the development of a specific aspect supporting quantitative and qualitative evaluation. For example Wakefield proposes a generalized operator action tree (OAT) representation and a modified HCR model to account for dependences between actions as well as a new approach to estimate probabilities of confusion matrix (CM) /Wa88a/. He describes a computer program HUMAN for performing the calculations corresponding to the modified HCR model.

2.3.6 An example of software supporting HRA

The iterative nature of a PRA means that the human failure event data base may have to be updated several times. The documentation and data management tasks of an HRA (Human Reliability Analysis) make up about 20-40% of the analysis, depending on the familiarity of the analyst with the plant being studied. Operator Reliability Calculation and Assessment (ORCA) has been developed to alleviate the routine tasks of Human Reliability Analysis (HRA) /Do88/.

ORCA's basic purpose is to document all human failure events in a PRA. The program was designed to have enough "intelligence" to ask the user for specific information and to work its way through an abbreviated taxonomy. Fig. 2.8 shows the basic logic path that ORCA follows. The user is first directed by menus to name the data base to work on and to name an event to add to, delete from, or edit in that data base. If the add or edit options are chosen, ORCA asks the user to describe the human failure event as a slip or a mistake. Depending on this choice, ORCA asks for information sufficient to calculate a screening value for the event, which is done automatically. If the user wants to input more information, or perform a detailed calculation, or otherwise produce a final occurrence probability, ORCA queries for the appropriate information and the final, mean probability and its 5th and 95th percentiles are automatically produced.

Once initial results are obtained and reviewed, any changes in parameters or key descriptors can be input by using the event editing mode. The results are automatically propagated through the ORCA data base and the HRA file is updated. Several report formats are available that have been found useful on past PRAs. ORCA also produces a summary listing of all events, one per line. ORCA is designed to be used on an IBM AT family of personal computers. ORCA was developed using Ashton-Tate's dBase III but is compiled so as to run without requiring dBase.

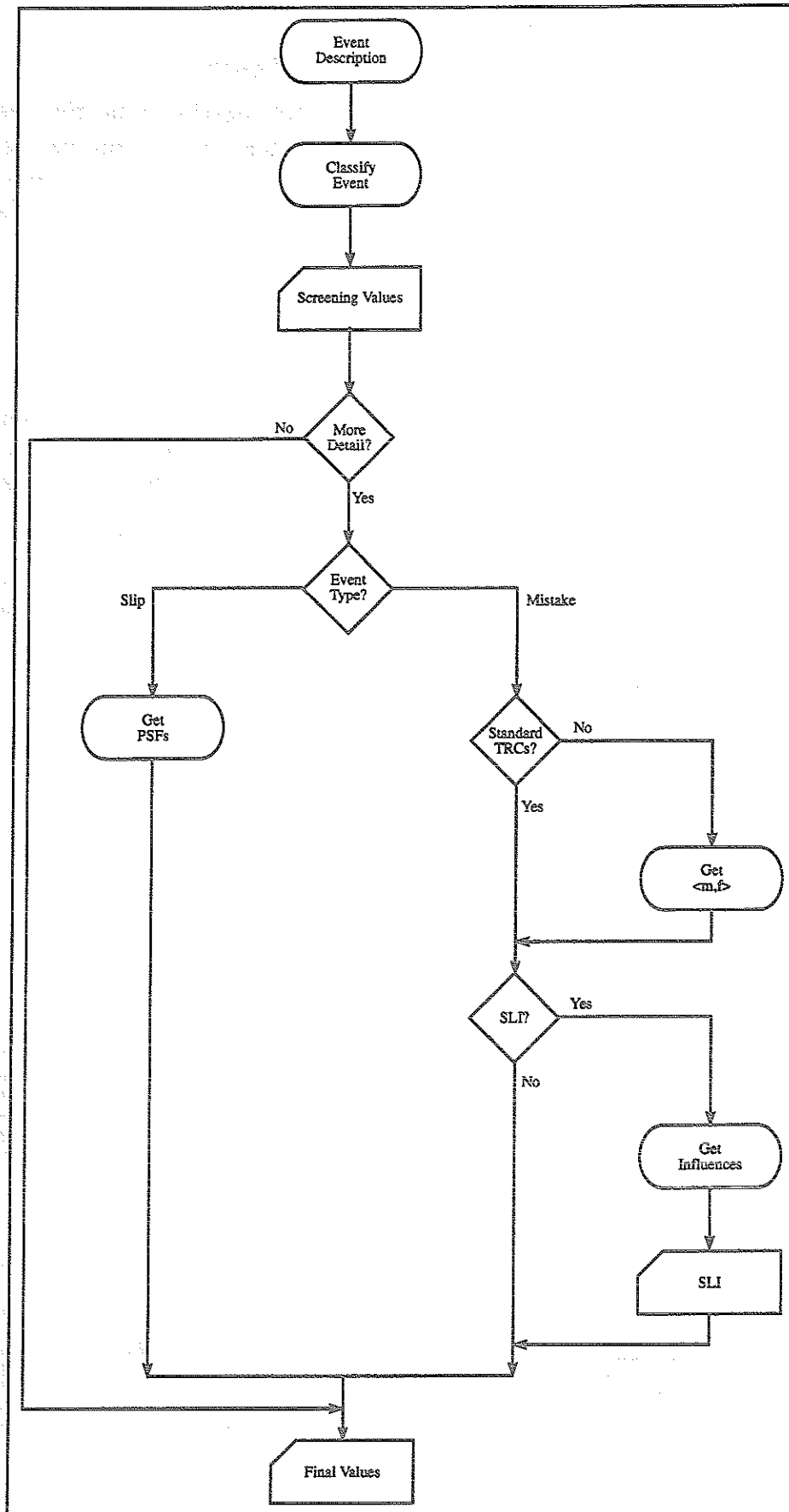


Fig. 2.8: ORCA logic structure /Do88/

2.4 Frameworks for incorporating HRA into PSA

2.4.1 Systematic Human Action Reliability Procedure (SHARP)

Systematic Human Action Reliability Procedure (SHARP), developed by Hannaman and Spurgin /Ha84/, is a framework for systematically incorporating human-hardware interactions into PSAs. In the SHARP approach the responsibility for incorporating human interactions is shared between the system analysts and the human-reliability analysts. The objective of the framework is to help in defining the types of interactions that are important to risk analysis and to enable the analysts to incorporate them into the system analysis task.

In SHARP five types of human interactions with the system are distinguished.

Type 1	Before an initiating event, plant personnel can affect availability and safety either by inadvertently disabling equipment, during testing or maintenance, or they can improve the availability of systems by restoring failed equipment through testing and maintenance.
Type 2	By committing some error, plant personnel can initiate an accident.
Type 3	By following procedures during the course of an accident, plant personnel can operate standby equipment that will terminate the accident.
Type 4	Plant personnel, attempting to follow procedures, can make a mistake that aggravates the situation or fails to terminate the accident.
Type 5	By improvising, plant personnel can restore and operate initially unavailable equipment to terminate an accident.

Tab. 2.6: Type of human interaction according to SHARP

This categorization was found to be useful for understanding how human interactions were treated in other PRA studies, apart from the five mentioned above.

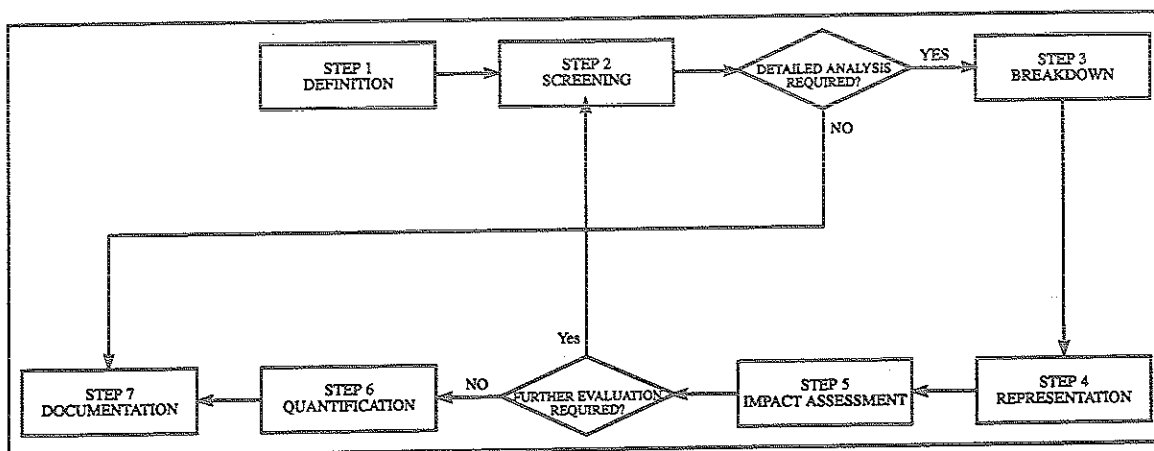


Fig. 2.9: Links between SHARP steps /Ha84/

The SHARP framework matured in the form of seven steps. Each SHARP step has defined objectives, inputs and outputs, activities and rules. The links between each step and the key decision points are shown in Fig. 2.9. The goals for each step are as follows:

1. Definition	To ensure that all different types of human interactions are adequately considered in the study.
2. Screening	To identify the human interactions that are significant to operation and safety of the plant.
3. Breakdown	To develop a detailed description of important interactions by defining the key influence factors necessary to complete the modelling. The human-interactions modelling consists of a representation (e.g., qualitative model), impact assessment, and quantification.
4. Representation	To select and apply techniques for modelling important human interactions in logic structures. Such methods help to identify additional significant human actions that might impact the system logic trees.
5. Impact Assessment	To explore the impact of significant human actions identified in the preceding step on the system logic trees.
6. Quantification	To apply appropriate data or other quantification methods to assign probabilities for the various interactions examined, determine sensitivities, and establish uncertainty ranges.
7. Documentation	To include all necessary information for the assessment to be traceable, understandable, and reproducible.

Tab. 2.7: SHARP steps

There are the following basic activities in these consecutive steps:

Step 1

The basic logic trees developed by the systems analysts from the functional description of the plant are enhanced in order to better and possibly fully describe human interactions, i.e. to ensure that all different types of human interactions are adequately considered in the study. To facilitate relevant analyses a classification scheme suitable for describing the types of human interactions modelled in PSA study (before, at and after initiating events), should be developed. Various sources of information (including emergency procedures, maintenance and test operations) should be considered.

Step 2

The logic trees, enriched with human interactions, are screened to select only the most important human interactions for further analysis. For that purpose a screening technique to rank and select key interactions for detailed analysis should be developed. The typical techniques are based on: (1) judgement (e.g. location in the tree logic), (2) coarse screening (e.g., by evaluation of the human impact in hardware cut-sets) or (3) fine screening (e.g., quantifying both generic hardware failures and human interactions based on behaviour type).

Step 3

Each key interaction is subdivided into tasks and subtasks associated with specific equipment or procedures, in order to define influence factors necessary for a complete modelling. The tasks/subtasks require enhanced descriptions by including the appropriate influence parameters (e.g., the behaviour type, method of detection, time allowed, alarms available, feedback for actions, coordination required, stress level, and other performance shaping factors) to establish the basis for selecting a model or representation of the human interaction.

Step 4

The key human interactions are explicitly modelled to include the possible options and a representation is selected (e.g., THREP, OAT, etc.) taking into account availability of data and/or expert opinion to support selected representation (if data not available, alternative representations can be considered). For selected representation a specific description is required (e.g., alternate choices or mistakes). In this way, it is possible to identify additional significant human actions that have an impact on the system logic trees.

Step 5

The system logic trees are re-modelled in order to explore the impact of significant human interactions (on initiating event, common-cause failures, reliability of systems, event trees logic, and sequence quantification). Then the candidates for re-screening and grouping are chosen. To analyse the impact of derived human interactions the screening techniques of Steps 1 and 2 can be applied. The human interactions are then grouped for quantification (according to defined criteria of grouping).

Step 6

The human actions identified in Steps 4 and 5 are quantified for incorporation into the PSA study accident-sequence quantification, including the evaluation of the uncertainties associated with the data. The most appropriate data base should be selected, which might include: time-reliability curves, published data or data obtained from experiments, human reliability models and opinions of experts.

Step 7

The results are documented in order to include all the necessary information for the assessment to be further used and applied. Documentation for each event tree should contain: data, assumptions, models, representations, list of important dependencies and sensitivities (associated with decision-making, times for actions, etc.) as well as a quantitative statement of the human impact on the core-melt frequency.

SHARP is considered to be a useful framework for incorporating various human reliability events into the equipment-oriented logic of PSA /Ca88, IA91c/. As was mentioned, probabilities of these events are evaluated using various human reliability quantification techniques or are taken from data bases if available.

2.4.2 Procedure for Conducting Human Reliability Analysis in Probabilistic Safety Assessment (IAEA's Procedure)

The Agency endorses the increased use of PSA and has established guidelines to facilitate the use of PSA in member countries /IA92a/. To support these PSA Guidelines a number of technical documents are being written to provide more complete and detailed technical

guidance in specific areas. A report /IA92b/ represents a technical document giving Procedure for Conducting Human Reliability Analysis (HRA) in Probabilistic Safety Assessment. The objectives of this report, according to intentions of its authors, are:

- to represent a practical and nearly standardized approach and terminology for HRA in PSA;
- to describe a framework in which different types of human actions are related to specific parts of a PSA;
- to describe acceptable methods and data sources for analysing human actions; and
- to show how the HRA is integrated into the rest of the PSA.

Human Interactions (Actions/Errors) and influences can have a variety of means of affecting safety or risk. It is important to be able to relate them to the PSA structure to understand their potential effect on risk or safety. Three categories can be defined that enable HRA studies to be compatible with PSA structure:

- Category A (Pre-initiators) - actions and errors that cause equipment/systems to be unavailable when required post-fault;
- Category B (Initiators) - actions and errors that either by themselves or in combination with equipment failures lead directly to initiating events/faults;
- Category C (Post-initiators) - actions and errors occurring post-fault. These can either be performing safety actions or actions/errors that aggravate the fault sequence. These interactions can be separated into three different types listed in Table 2.4.

The main components of an HRA study are as follows /IA92b/:

I	Task definition / understanding	A precise definition of the overall task (or error) being addressed. This will provide the boundaries for the study and will match the interface requirements of any assessments being performed.
II	Task analysis	This is the detailed process of understanding how individual elements fit together in order to successfully accomplish the overall task. Essentially it is an information collection and task logic structure process. There are a wide variety of formal task analysis methods available but irrespective of any specific method chosen the HRA analyst (for PSA) must ensure he/she has gathered and structured sufficient information to have a clear appreciation of the individual steps needed to achieve the overall task.
III	Error identification	The error identification process comes directly from the task analysis. The HRA analyst must identify potential errors at each individual step of the overall task. In particular the potential for mistakes should be noted.
IV	Error recovery potential and mechanisms	The potential and mechanisms for recovering from the identified errors should be determined or judged. This will often require additional information that was collected during the initial task analysis.
V	Identify additional constraints	"Performance Shaping Factors" (PFSs), e.g. experience, level of training, stress level. All factors judged to be important for successful completion of the overall task should be identified.
VI	Quantification	Quantification of the success/failure of the overall task is based on using the information obtained in steps I to V to quantify the individual tasks or errors and their possible recovery routes.

Tab. 2.8: Main Components of an HRA study /IA92b/

Steps I to V can be termed the qualitative assessment phase and step VI is the quantitative phase.

The process of incorporating HRA into PSA can be separated into seven basic tasks. Each task consists of inputs, activities and outputs. For each category of events (A, B and C) the tasks required have the same form. The tasks and their goals are as follows:

1. Definition	To ensure that all candidate human interactions of the different categories are adequately considered in the study.
2. Screening	To identify the human interactions that are significant to the operation and safety of the plant and to avoid wasting analysis resources on the others.
3. Qualitative Analysis	To develop a detailed description of important human interactions by defining the key influence factors necessary to complete the modelling.
4. Representation	To select and apply techniques for depicting important human interactions in logic structures.
5. Model Integration	To describe how the significant human actions are integrated into the plant and the system models of the PSA.
6. Quantification	To apply appropriate data or other quantification methods to assign probabilities for the various interactions examined, determine sensitivities and establish uncertainty ranges.
7. Documentation	To include all necessary information for the assessment to be traceable, understandable and reproducible.

Tab. 2.9: Incorporating HRA into PSA

The extent and character of the HRA will inevitably depend quite strongly on the ultimate objectives of the PSA. The effective integration of the HRA elements into the PSA requires knowledge of both HRA and PSA. Ideally a human reliability analyst with an understanding of basic PSA methods and Nuclear Power Plant (NPP) operations would be used to co-ordinate the HRA integration. Where this is not possible then a systems analyst from the PSA team should be assigned to co-ordinate and lead the HRA studies with additional support as necessary from other systems analysts. Close contact with the station staff is essential to ensure that both the PSA model and HRA evaluations within it reflect actual station operations and practices. If the PSA is being performed at the design stage then the HRA leader should consult knowledgeable staff from other operating stations to ensure that assumptions about operating practices, procedures and training are reasonable.

2.5 Approaches based on integrated simulation or AI methods

2.5.1 Cognitive Environment Simulation (CES)

The Cognitive Environment Simulation (CES) concept /Wo86, Wo87/ has been worked out within a three-phase research programme sponsored by the U.S. Nuclear Regulatory Commission. The aim of research work was to develop and validate methods for modelling the cognitive behaviour of the nuclear power plant (NPP) personnel during emergency operations. In the second phase of this work a model related to operators' intentions to act in NPP emergency situations was developed using an artificial intelligence (AI) technique. This model is a basis for performing Cognitive Environment Simulation (CES). The system is an analytic computer simulation tool /Wo87/.

The system is theoretically based on several scientific disciplines such as cognitive sciences or mind sciences and includes a variety of fields such as cognitive psychology and artificial intelligence. Models developed were called cognitive models. A methodology has been also worked out for evaluation of the human contribution to risk in probabilistic risk assessment (PRA) studies, named Cognitive RELiability Assessment TEchnique (CREATE).

As described in reports /Wo87, Wo88/ CES allows exploration of plausible human responses in different emergency situations. It can be used to identify what are difficult problem-solving situations, given the available problem-solving resources (e.g., specific procedural guidance, operator knowledge, man-machine interfaces). It provides the capability to establish how people are likely to respond. One can investigate e.g. what situations and factors lead to intention failures, the form of intention failure, the consequences of an intention failure including: errors of omission, errors of commission or common mode failures. CES might also identify situations where operators are likely to detect and recover failed systems or to correct an erroneous human action. These potential recovery actions can then be incorporated into the event/fault tree analysis.

Within CES some variables are called Performance Adjustment Factors (PAFs), which are similar to traditional Performance Shaping Factors (e.g., experience level, stress, organizational climate). PAFs provide the possibility for a human analyst to establish parallels between the cognitive environment simulated in CES and NPP situations to be investigated (Fig. 2.10). CES should also provide a tool for investigating the effects of changes in NPP person-machine systems including instrumentation, computer-based displays, operator decision aids, procedure changes, training, multi-person or multi-facility (e.g., technical support centre) problem solving styles.

CREATE (Fig. 2.11) was developed to specify how CES can be used to enhance the measurement of the human contribution to risk in PRA studies. In CREATE, CES is run on multiple variants of accident sequences of interest. The variants are selected to represent parametric combinations of a plausible range of values along the dimensions that contribute to cognitive task complexity. The goal is to identify sets of the minimum necessary and sufficient conditions (characteristics of the situation and/or the operator) that combine to produce intention failures with significant risk consequences. Once the range of plausible intention errors and the conditions under which they arise are identified, a quantification procedure can be initiated to assess the likelihood of these intention errors.

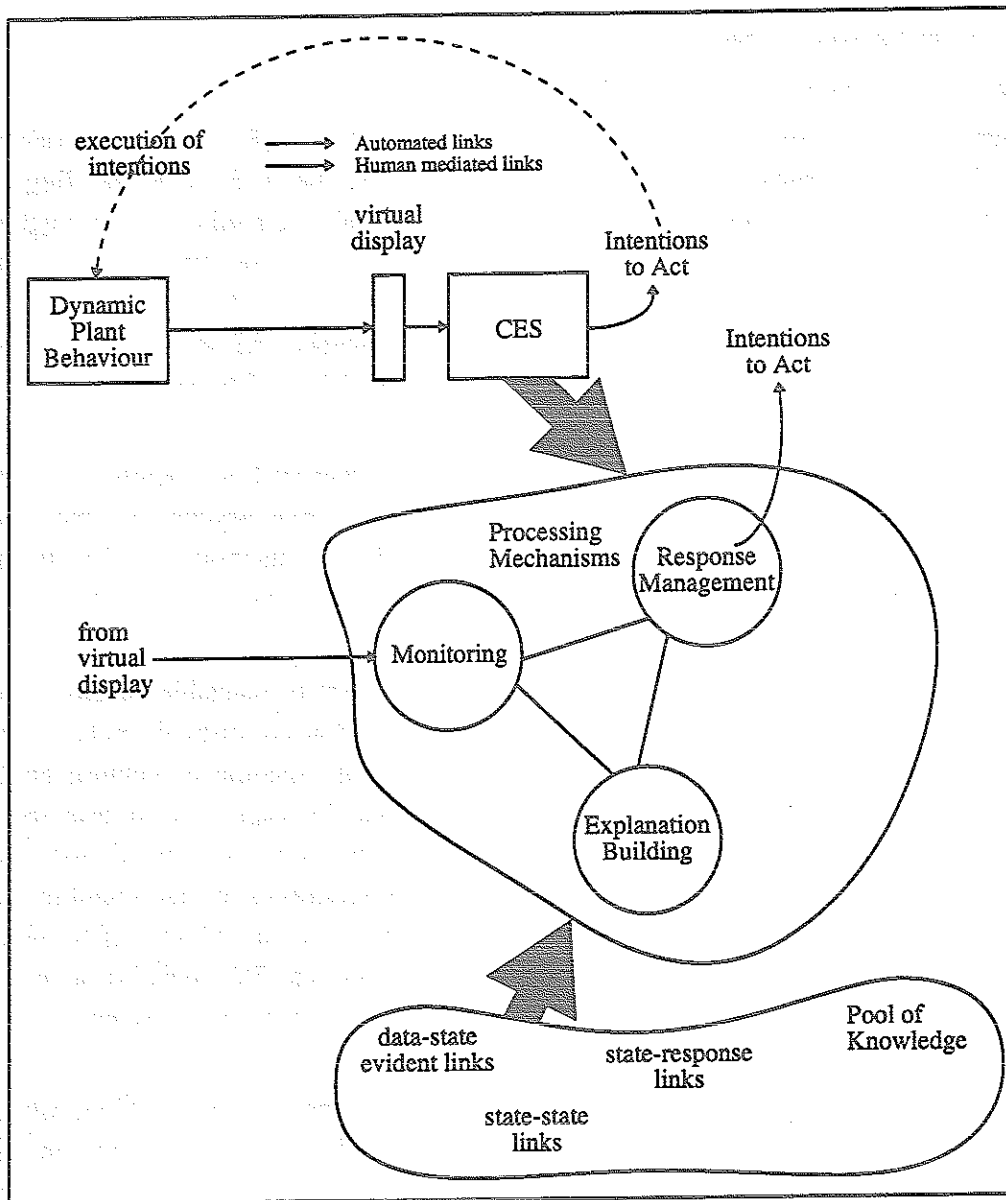


Fig. 2.10: Different kinds of processing in CES /Wo87/

The methodology involves two main stages:

- (1) A modelling stage where CES is used to find situations that can lead to intention failures and therefore to erroneous actions; and
- (2) A systems analysis input stage where the results of the cognitive modelling are integrated into the overall systems analysis.

The main steps in the modelling stage are as follows:

- decide what NPP situations to investigate with CES and how these situations map into the simulation environment,
- set up CES to be able to run NPP situations,
- run CES over a plausible range of demand and resource settings, given the analysis of this plant,
- to identify the conditions which produce intention failures and the actions that follow from an intention failure.

The main steps in the systems analysis input stage are:

- modify the event/fault trees to reflect the effects of intention errors identified in the modelling stage,
- employ a quantification procedure to assess the likelihood of these intention errors,
- combine the intention error with execution error estimates.

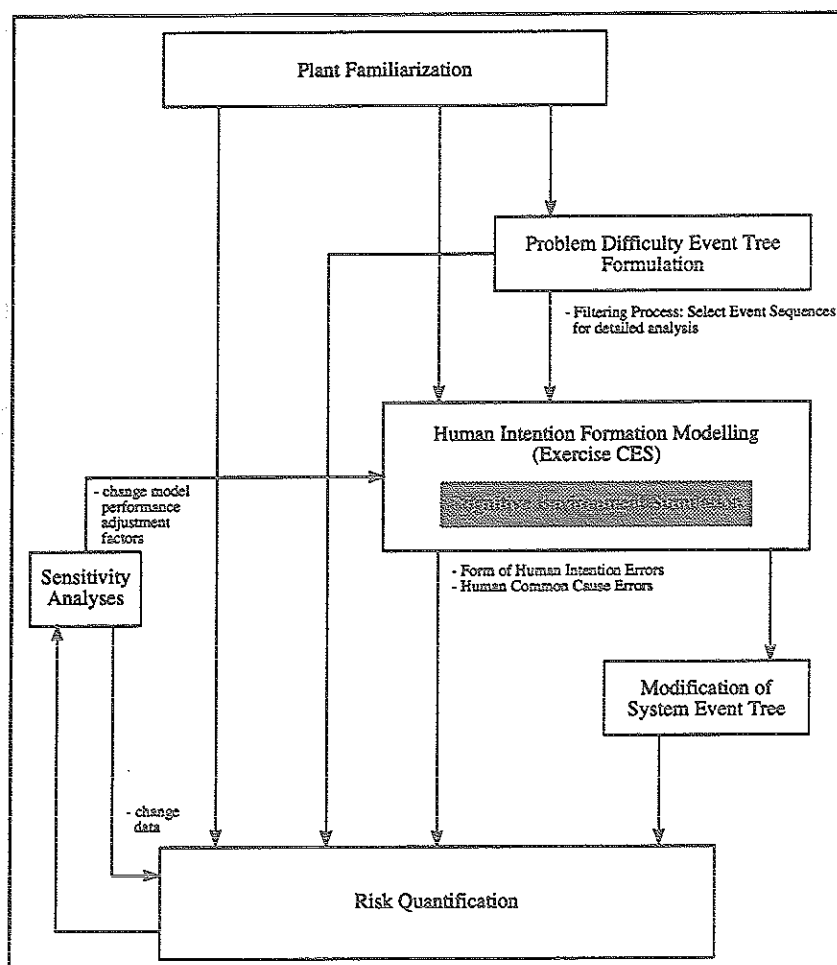


Fig. 2.11: Cognitive Reliability Assessment Technique (CREATE) /Wo87/

The results of the CREATE modelling stage will affect overall systems reliability analysis in two ways: (1) it will lead to modifications of the event/fault trees and (2) it will provide quantitative estimates of human reliability as input to overall systems reliability frequency estimates. In the absence of empirical data, expert judgement is, however, required. It is proposed to use techniques available for eliciting frequency judgements from groups of experts (e.g., technique used in Reactor Risk Document /Nu87/ or in the SLIM-MAUD procedure /Em84/).

Overall human reliability estimates can be obtained by combining the probability of human intention errors with the probability of human execution errors. It is proposed to calculate the joint probability of a failure as the sum of these probabilities (assuming that intention errors and execution errors are independent). Execution error probabilities might be generated using the standard HRA estimation techniques (e.g., Swain & Guttman or the SLIM-MAUD procedure).

CES is a specific instance of an artificial intelligence problem solving system EAGOL. The EAGOL architecture embodies unique capabilities for reasoning in dynamic situations that include the possibility of multiple faults. EAGOL is a software system and proprietary product of Seer Systems which builds on the conceptual framework of the CADUCEUS AI problem-solving system developed for medical problem-solving applications.

CES currently runs on Symbolics-3600 class computers and is planned to run on Sun-3 class computer as well. It is built on top of the EAGOL AI software shell which uses CommonLisp and the Flavors object oriented programming environment. The basic software engine and knowledge base is transportable to several machines that can support the CommonLisp/Flavors software environment such as Symbolics, Sun, or Vax computers.

Remarks:

The authors of CES are aware of the complexity of the proposed simulations. They assume that cognitively challenging conditions represent classes of NPP situations. While CES may run on only one or a few incidents that are instances of the class, it is necessary to aggregate the probabilities of all the ways this cognitively challenging situation could arise in estimating the likelihood of this contributor to intention failures.

They propose, in cases of the absence of empirical data, to use the multi-expert judgement techniques, e.g. SLIM-MAUD procedure. It was not described how to take advantage of very complex simulations and use CREATE methodology to support such judgements. It seems that the entire methodology is not easily applicable to support the engineering HRA/PSA studies.

The authors outlined the next steps in the system development which are needed to take advantage of the capabilities of the CES cognitive model and the CREATE methodology:

- empirically validate the correspondence between CES and human behaviour,
- evolve the model's capabilities and its accessibility to the potential user community,
- further develop and refine the CREATE methodology through by exercises with cases of relevance to PRA.

2.5.2 An integrated modelling of operator cognitive behaviour

An integrated approach for modelling human and machine has been proposed at JRC-Ispra /Am87, Be88, Ca88/. The simulation of man-machine interactions is based on the System Response Analyser (SRA). This tool enables the analysis of nuclear plant subsystems in abnormal or emergency conditions focusing on the role played by the operator. The DYLAM (DYnamic Logical Analytical Methodology) technique is used to drive the entire simulation. The second version of the this technique, DYLAM-2, has been developed to cope with the more general cases of multiple occurrences (states of components) during the simulation process.

Having generated a 'mother' sequence (that is a combination of initial states), DYLAM-2 is able to generate further 'descendent' sequences which can be provoked by possible random-in-time changes of states of system components. The generation law of the forest of descendants is described by failure and repair rates for independent events and/or by conditional probability matrices for dependencies among 'states' and physical variables (e.g. pressure, temperature, etc.).

The analyst can choose different modes of generating sequences:

- a systematic mode, in which mother and descendent sequences are generated sequentially up to a desired resolution level according to the heuristic rules built into the code;
- an interactive way, in which the analyst prefixes the initial conditions, that is he selects an initial mother sequence and then lets DYLAM generate its possible descendants;
- a pseudo-Monte Carlo procedure in which many failure occurrences are forced, as they could hardly be simulated otherwise because of their low probability of occurrence.

DYLAM has been designed not only as probability estimator, but also for exploring the space of the 'possible' events; in this case the probability attributes are utilized only to guide the event sequence procedure generation. The system permits the analyst to model dependencies which may be identified both in component and human models.

Due to the complexity of the control of a nuclear power station, plant modelling has been concentrated on processes within the primary circuit and main automatic control mechanisms in order to allow sufficient extension of the human model. In principle the intention was to focus on the study of accidents and thus on those parameters which are of main interest for the operator in such circumstances. For this reason the safety parameter display panel (EPRI Project, 1982) concept is taken as a reference and the simulation of the plant is centred on the evaluation of these parameters.

During research work a new model for very rapidly simulating the plant response became attractive for the development, based on the response surface technique. For the effective evaluation of the transient behaviour of the most significant parameters the model of the plant is to be approximated by polynomial expressions. Data for a variety of configurations and accident scenarios for a PWR are to be prepared for testing (among these: station blackout, small loss of coolant, steam generator tube rupture, secondary circuit line break, etc.).

The framework within which complementary aspects of human behaviour are being modelled includes three types of cognitive processes (Fig. 2.12):

- (i) High-level decision making (HLDM), such as situation assessment, planning, plant modification etc.
- (ii) Low-level decision making (LLDM), whereby the actual relationship amongst intention-execution plant response is implemented.
- (iii) Errors, which account for slips or mistakes made at any level of the human activity, depending on a variety of factors, including environmental conditions, internal beliefs, fixations, stress etc.

The human model under development is framed within a structure that has the models of the machine and of the operator as interactive counterparts /Am87/, while DYLAM acts as a driver and controller of the entire system. Within this frame, the assessment of a situation (diagnosis) and the formulation of a strategy (planning) are considered as a High Level Decision Making (HLDM) process, because they imply long-term planning as well as the analysis of the plant as a whole and possibly reasoning about the evolution of physical phenomena. Formulated at this level the strategy of actions and interventions is implemented in the simulation process using the Low Level Decision Making (LLDM) model.

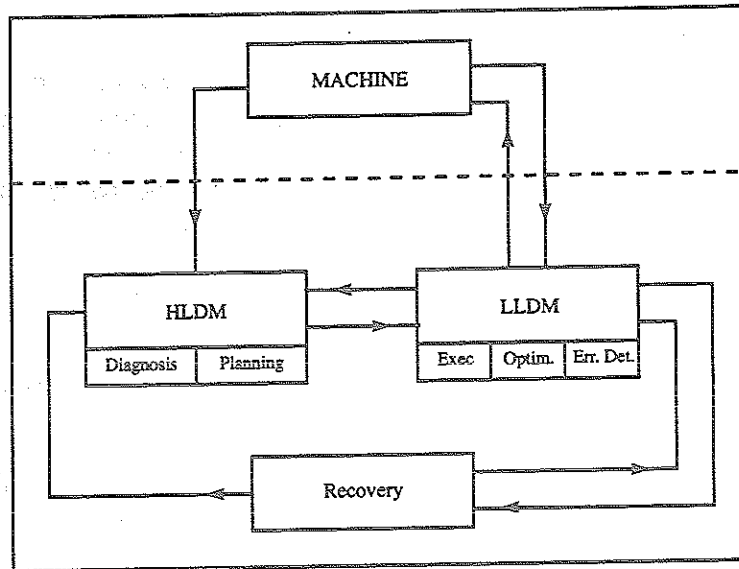


Fig. 2.12: General framework of the human behaviour model /Am87/

Two basic assumptions were made:

- (1) The operator behaviour develops within a goal-oriented hierarchical structure.
- (2) Fuzzy-logic is the formalism to handle imprecise and uncertain understanding of the reality.

Errors can be made at any time of the sequence under study or at any level of the human behaviour model. Thus an error mechanism is continuously active and can affect diagnosis, planning or plan execution. For this reasons the error process is not treated as a human activity as such but it constitutes the object of a separate model, which runs in parallel to the rest of the simulation. Recovery of a given situation is also modelled in the human activities flow chart: recovery can be triggered by the detection of some anomalies or recognition of slips or mistakes. The process of recovery expands over low level as well as high level decision making and it refers not only to human errors but also to additional system failures occurring during the incident.

The fundamental steps of the HLDM model contain the processes of diagnosis and planning. Diagnosis is rarely a point-type process, it entails a continuous interaction with planning and verification. A preliminary model of diagnosis formulation has been created which is based on a tree like structure. As soon as a minimum number of data has been gathered, the model is capable of working its way up the tree to the formulation of a first diagnosis. This activation of a set of hypothetical schemata relies essentially on some fundamental cognitive primitives such as: similarity-matching and frequency-gambling. Then the structure is re-examined in a deductive way (top-down), more in a planning perspective, with the goal of verifying the first diagnosis, which was based on a limited and possibly incomplete set of data.

The successive planning process leads to the identification and selection of actions or procedures. Starting with a main intention or 'top goal', the result of this process is the construction of a hierarchical network, where the different sub-goals and sub-tasks to be performed by the operator are schematically arranged in a 'tree' type structure and are linked to each other by different gates such as 'AND' and 'OR' (Fig. 2.13). A 'goal' is an element of the network at any level. An 'act' is the last elementary goal of the network, i.e. the action that the

operator will perform (close a valve, stop a pump, etc.). A 'task' is the set of acts the operator has to perform in order to attain a certain goal at any level.

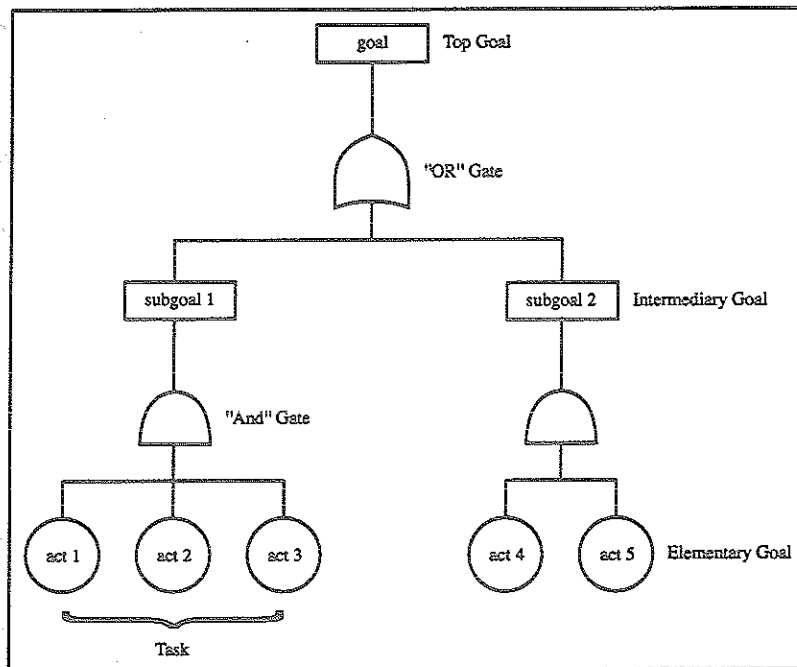


Fig. 2.13: Goal, task and act concept /Am87/

In the approach proposed the tree structure is, in most cases, built by a network of precompiled 'a priori' goals and sub-goals. They represent what is contained in procedure manuals or what constitutes the operator's familiar and pre-experienced procedural sequences. However, when the first diagnostic process does not lead to a set of pre-compiled goals, the construction of the strategic tree results from the definition of each goal in terms of its preconditions and effects. In this case, in order to develop an automatic construction of the tree, reference has been made to AI studies, where the planning process has been classically supported either by two domain-independent knowledge bases: a model of the world and a set of goals to modify this world, or different levels of pre-compiled plans /Am87/.

The high level planning level proposed departs from traditional AI models due to some new considerations:

- (A) An integrated view is considered, where diagnosis interacts continuously with planning either in activating new frames containing new intentions to satisfy, or in ceaselessly updating the world model, entailing the development of new strategies, even for the same goal. Moreover, interacting with execution causes the installation, in the tree, of OR gates, which allow the operator to select among alternatives to attain a goal, if the first chosen route has failed.
- (B) An uncertain and imprecise definition of the goals is allowed, entailing the strategy development to rely on fuzzy instead of classical logic.
- (C) The construction of the tree as a product of the planning process is achieved by recurrent exploration of three knowledge bases (KB):
 - (1) A KB containing the description of all goals that the operator may attain, in terms of attributes that identify the boundary of a goal such as pre-conditions, expectancy, side-effects (these three attributes can be defined in crisp or fuzzy terms), upper and lower

goals, etc., and parameters governing the unfolding of the goal execution (priority, certainty, satisfaction).

- (2) A KB containing the structural operator representation of the plant in terms of systems, sub-systems and components.
- (3) A KB containing the operator representation of the process, based on a qualitative and symbolic approach as, for instance naive physics or fuzzy models /Am87/.

The HLDM process terminates, i.e. the goal tree structure is completed, when a level composed of all elementary goals has been reached. Then the LLDM model carries over the actual execution of the developed strategy. The LLDM model interacts with the simulation of the plant and its interfaces. Navigation through the tree is regulated by the values assumed by the parameters of the goals which are:

- (1) the degree of priority (GDP), which expresses the measure of hierarchical sequence between goals;
- (2) the degree of membership (GDM), which defines the measure of the dependency between a goal and its directly superior goals;
- (3) the degree of satisfaction (GDS), which represents the correlation between the result of a goal being performed and the operator's expectation;
- (4) the degree of certainty (GDC), which represents the measure of the attainment of a goal as a function of GDS and GDM.

The hardware architecture on which the System Response Analyser (SRA) is being developed is based on a SYMBOLICS, LISP machine, and a network of SUN stations. The human model is represented in LISP on SYMBOLICS computer, because this language was considered to be suitable for simulating the processes of decision making and execution according to prestructured plans. Implementation of the software package KEE (Knowledge Engineering Environment) has been planned which would allow the easy formulation of structures representing the operator knowledge bases. The implementation of the machine model was planned to take place on a SUN work station. The DYLAN simulation programme has already been implemented on this work station /Am87/.

Remarks:

The approach described above was not primarily developed for the quantification of human reliability. The overall scope of the project was intended to support analyses of man-machine interactions in accident conditions in order to: (1) Improve the system design; (2) Improve operator training; and (3) Improve the support given to the operator in case of decisions. The outcome of these simulations, as the authors propose, would be processed by an expert system that would be able to recognize novel information, to chain this information to already existing data bases and eventually to organize its rules to be supplied to the operator in a decision support system (DSS).

2.5.3 Cognitive Simulation Model (COSIMO)

The model named COSIMO (COgnitive SIMulation MOdel) has been developed for modelling a human operator controlling a complex environment /Ca89/. In the model the execution of "lower" level activities, such as recognizing, diagnosing, planning etc. (as opposed to "higher" cognitive activities such as intentionality, parallel reasoning, adaptive behaviour etc.), is guided by "meta"-cognitive processes. Both these types of activities are related to psychological

theories and are present in the simulation: the selection and ordering of processes depends on the evolution of the man-machine interaction itself and a number of data relative to the meta-processes, describing the state of the operator at the time and during the sequence of events.

In this framework several models of activities at the "lower" level are identified. They are namely:

- cognitive filter, which selects data on the basis of their salience;
- diagnosis, which chooses an explanation;
- confirmation, which verifies hypotheses or expectations;
- planning, which defines a sequence of actions to be executed.

At the psychological level, these activities are carried over exploiting the basic mechanisms of Similarity Matching (SM) and Frequency Gambling (FG). The mathematical forms that allow a semantic interface to be created between the system simulation and the operator cognitive model is based on fuzzy logic, which is a theory well suited for representing the approximate operator's knowledge and the imprecise understanding of reality.

The metaknowledge mechanisms account for the key peculiarities of complex systems like a human being, such as dynamicity and non-linearity of the interaction process, self-referentiality of the human being, effect and trust of the presence of other human beings and the interface mechanisms for contributing to decision making. These mechanisms allow resetting, rescheduling or maintaining the current cognitive activity. However, the development of the models of meta-cognition is at its preliminary stages and, for the time being, only some simple rules have been implemented for testing purposes.

The architecture comprises two fundamental components: a virtually limitless Knowledge Base (KB), containing both declarative and procedural knowledge structures, represented as frames, and a limited, serial but computationally powerful Working Memory (WM). The conceptual architecture of the COSIMO model is illustrated in Fig. 2.14.

The content of KB is structured in frames or schemata which represent operator knowledge about the plant in the form of geometrical structures (connections, locations, states of components), process representations (variables behaviour, causal relations, functional landmarks and thresholds), and control sequences (actions, tasks, procedures and respective effects). Two main types of frames are distinguished: knowledge frames (describe only the processes and the structure of the system in terms of general physical and engineering principles, as well as rules of thumb) and rule frames (encompass pre-definite plans of actions for different situations).

WM represents a working area for the cognitive processes, the management and the temporary storage of data. The mechanisms for bringing products of stored knowledge units into the WM are: SM between calling conditions and attributes of KB schemata; and FG for selecting a final candidate among a number of partially matched candidates, on the bases of frequency and recency of encounter.

The Blackboard (BB) framework is proposed as a structure to represent operator reasoning and decision making. The Blackboard structure consists of different levels of abstraction which correspond to the various kinds of structured objects that may be dealt with by the model. The Blackboard also consists of the Agents (Knowledge sources) and the controller, governing the agents' actions. Two different BBs can coexist and interact with each other: a Domain-BB (DBB) and a Control-BB (CBB). In particular, the CBB generates tasks which are added to the tasks of the DBB.

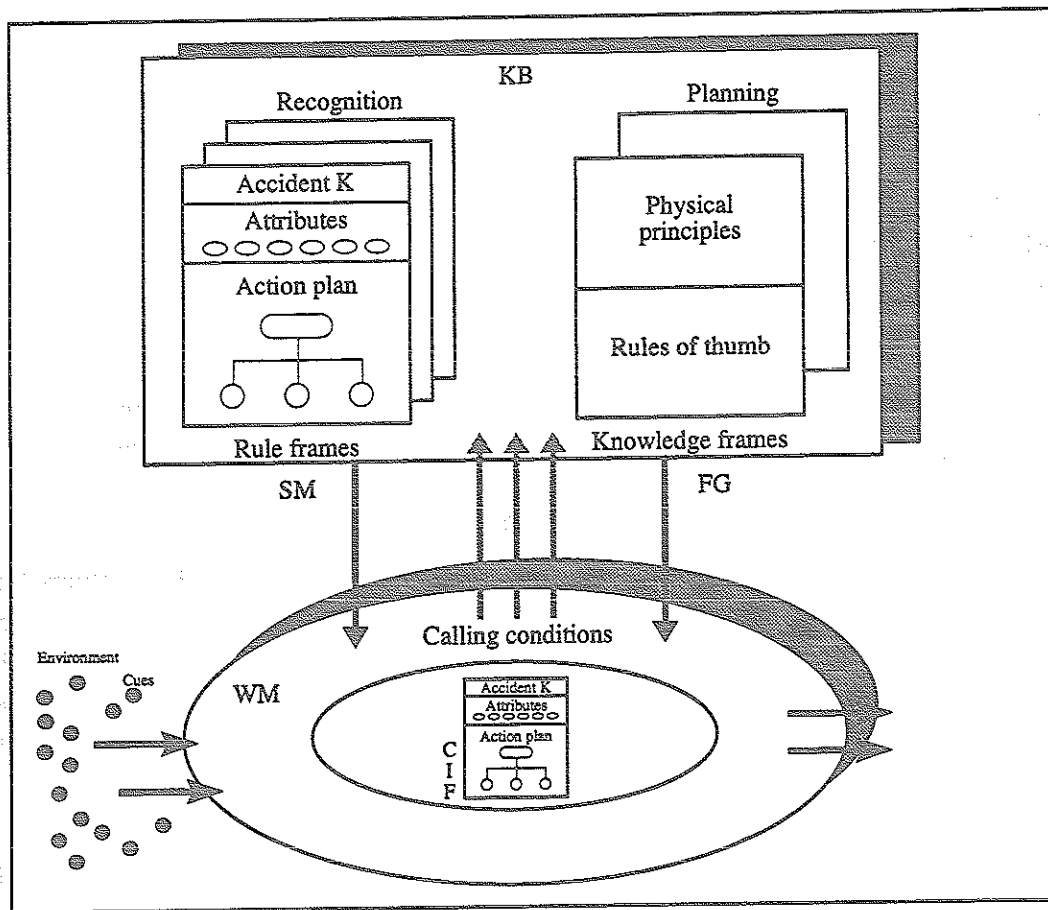


Fig. 2.14: Conceptual architecture of the COSIMO model /Ca89/

This framework enables the system to tackle dynamic situations, where changes of lines of reasoning and changes of control strategies are needed to cope with the intrinsic non-monotony of the problem. It allows the representation of the adaptive behaviour of the operator. Fig. 2.15 represents the basic BB cycle which applies each time the Controller selects a new task to execute.

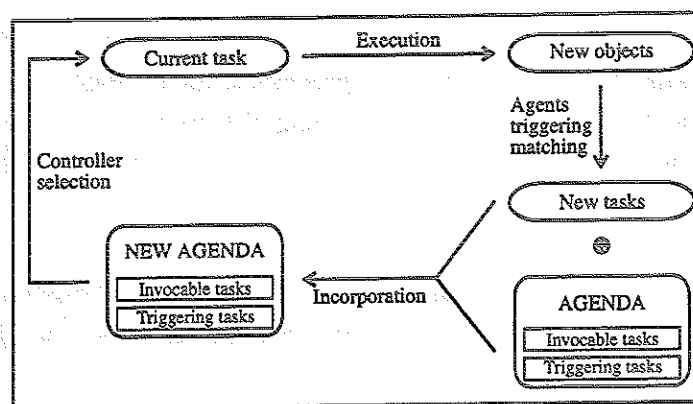


Fig.2.15: Blackboard basic cycle /Ca89/

In the model four correlated cognitive activities are simulated: data acquisition, diagnosis, confirmation and planning execution. These activities are defined by means of their related levels and agents at the DBB level. The interactions of these activities, i.e. their interruption, chaining, shifting, is governed by the CCB.

Remarks:

The important peculiarity of the model is its ability to simulate the evolution of a human-machine interaction which is adaptive both to the internal state of the operator and to the history of the human-machine system. The flexibility of the simulation is assured by the blackboard architecture. In the model the synergism of AI and psychology approaches was exploited.

The model developed relies, however, on strong theoretical assumptions, and therefore makes more demands on validation. The validation should be understood as a guide in the process of building the model: data collection, even in the real world or laboratory, can continuously provide requirements for the advances of modelling. The validation phase represents, thus, the future necessary step in the development of the model. An alternative would be to start from data collected in real situations and to formalize a model on this basis.

2.6 Remarks on available HRA techniques

2.6.1 Evaluation of quantitative HRA methods used in PSA

2.6.1.1 Human Cognitive Reliability (HCR) model

The HCR method has been developed for the estimation of the non-response probability of an operating crew. The approach can be used for time-constrained emergency situations, that is for diagnosis and decision making tasks.

The numerical accuracy and consistency of the time correlations used in the HCR model for skill-, rule- and knowledge-based behaviour are not sufficiently quantified, especially probabilities lower than 10^{-2} are difficult to validate. However, HCR is one of the few methods which is not based on expert judgement only, but also on simulator data.

The HCR method is easy to use, and analysts do not need excessive training. The acceptability of the method is good. HCR provides an easy utilization of simulator experiments for PSA purposes.

2.6.1.2 Technique for Human Error Rate Prediction (THERP)

THERP is the most widely used technique to date and allows the assessment of pre-accident and post-accident operator actions. The THERP Handbook offers a data base of human error probabilities and several models to consider performance shaping factors like stress or personal redundancy. The Handbook also offers a model for calculating uncertainty bounds. Most 'hard' data contained in the 'Handbook' are typical of skill- and rule-based behaviour. Therefore THERP is considered to be particularly effective in quantifying slips and rule-based mistakes (see Table 3.2). The diagnostic models presented in the Handbook to assess knowledge-based actions are generated only by expert judgement.

Analysts need some training for correct application of the THERP method. The acceptability of THERP is very high; the THERP methodology was used in almost every PRA published in recent times.

The ASEP method is a short version of the THERP Handbook. ASEP offers a step by step procedure to assess operator tasks. In comparison to THERP the application of ASEP produces more conservative results.

2.6.1.3 Human Error Assessment and Reduction Technique (HEART)

HEART calculates human error probabilities (HEP) as a function of performance shaping factors (PSF) identified for a particular task.

More than 30 PSF's can be considered for quantifying HEPs. It is possible to estimate all kind of operator action. HEART presents uncertainty bounds for every HEP.

Certainly HEART offers a lot of parameters to assess operator actions, but the underlying data base is as yet unpublished.

The method is relatively easy to use, and no great effort is needed. This relatively new technique has been applied several times for PSA purposes.

2.6.1.4 INTENT

INTENT was developed particularly for the assessment of decision making errors caused by incorrect intentions for taking actions. To this end 20 basic error probabilities are presented which can be modified using 11 performance shaping factors.

The application of INTENT is not easy due to the insufficient definition of basic error probabilities and performance shaping factors. However, INTENT is basically qualified for the quantification of decision making errors. In the frame of PSA, INTENT should be combined with other HRA models.

2.6.1.5 Success Likelihood Index Method (SLIM)

The basic premise of SLIM is that human error probabilities depend on combined effects of PSF's. Therefore several influence factors can be taken into account. SLIM can be used to assess test and maintenance operator tasks and provides sensitivity analyses to identify weak points in ergonomic plant design and procedures. The method comprises two modules: SLIM-MAUD and SLIM-SARAH (see 2.3.4.1).

Uncertainty bounds cannot be obtained with SLIM and have to be derived by using other methods. The application of SLIM requires well-known probabilities of reference tasks which are used as calibration factors, but in some cases these factors are difficult to obtain, and analysts have to rely on estimations. Especially if they use the MAUD software, analysts will need a good deal of training to apply the method effectively.

SLIM has been used in several PSA/PRA's and also in the course of the ISPRA Human Factor Benchmark Exercise. This indicates a high acceptability.

2.6.1.6 Socio-Technical Approach to Assessing Human Reliability (STahr)

The STahr method has a special capability for sensitivity analysis and is basically applicable to a wide range of situations (pre-accident and post-accident tasks, as well as skill-, rule-, and knowledge-based operator behaviour).

Using the STahr method it is possible to consider dependencies between different PSFs, but no consistent theory exists for this purpose, and in this sense the STahr model is speculative, and the calculation of human error probabilities requires theoretical justification and mathematical evaluation.

The STahr method depends strongly on subjective opinions. The application of STahr requires much time and personnel efforts. The acceptability is not very high, the method has only been applied in a few safety analyses.

2.6.1.7 Paired Comparison (PC) technique

The Paired Comparison method can be used to assess all types of operator actions. The validity depends mainly on the knowledge of the experts who are asked to compare a set of pairs of tasks and to decide which of them has the higher likelihood of error. Two or more tasks with well-known error probabilities are used to calibrate this scaling. But sometimes it is difficult to find suitable calibration data, and so the PC method is relatively subjective. Apart from this the

PC technique is easy to use, and normally experts do not require extensive training. PC is a well established method and has been applied in some PSA/PRA studies.

2.6.1.8 Absolute Probability Judgement (APJ)

APJ is one of the most extensively developed methods in assessing human error probabilities. Its advantages are primarily the ease of use (training as in using computerized models is not needed), the wide range of situations (pre-accident as well as post-accident tasks, skill-, rule-, and knowledge-based behaviour) which can be considered, and applicability onto misdiagnosis which makes a situation worse.

Using APJ as an HRA model is time-intensive and requires experts who are familiar with the area of interest and probability calculus. In order to compensate subjectivity more than one expert should be involved (e.g. using one of the known group methods) and in case of need an appropriate facilitator should be chosen.

As a sensitivity analysis is not possible, APJ cannot be used effectively in error reduction. Nevertheless the acceptability of APJ has to be classified as moderate. This statement is buttressed by the numbers of assessments done by APJ worldwide.

2.6.1.9 Confusion Matrix (CM) technique

The Confusion Matrix method concentrates on the aspect of misdiagnosing an accident situation. Thus only knowledge-based behaviour can be assessed. Additionally the method provides proposals for improvements in ergonomic plant design. No formal procedure for considering influence factors is included. As with other expert judgement methods the CM technique is based only on the knowledge and experience of the analysts. The general idea of CM is sound, but the utilization of expert judgements is based only on assumptions. Application of the CM technique is rather complicated, especially if analysts want to derive quantitative error probabilities. Therefore sometimes the method is used for qualitative purposes only.

2.6.1.10 Summary

The summary of evaluations is presented in Table 2.10. The criteria used in the Table should be understood as follows:

- Usefulness / completeness
 - assessment of different types of operator behaviour
 - assessment of different types of operator tasks
 - qualitative and quantitative results
 - insights for recommendations
- Validity / Accuracy
 - kind of data
- Ease of Use
 - special training of analysts
 - necessary equipment and time
 - necessary data requirements
- Acceptability
 - frequency of application

	Usefulness Completeness	Validity Accuracy	Ease of use	Accepta- bility
HCR	+	(+)	+	+
THERP	++	(+)	(+)	++
HEART	+	(-)	(+)	+
INTENT	-	-	-	(+)
SLIM	++	-	-	+
STahr	++	-	-	-
PC	+	-	-	+
APJ	(+)	(+)	+	+
CM	-	-	-	(-)

Tab. 2.10: Quantitative HRA methods and criteria used for evaluation

++ high; + moderate; - low

HRA techniques currently used in practice have proved very useful for various decision purposes during the design and operation of complex technological systems, although some of them represent the behaviouristic view of human operator behaviour. However, many misleading applications have also been encountered /Do90/. HRA and relevant reliability estimates depend heavily in most cases on expert opinions. The evaluation is often not properly documented as regards reasoning and attributes influencing human behaviour and therefore cannot be easily scrutinized. Thus following issues should be taken into account:

- limitations of using the currently available HRA techniques in practice;
- selecting scheme for application of HRA techniques and performing analyses by experts to cover various cases with regard to experimental and field data if available;
- selecting and training of experts; quality assurance aspects using single expert analysis or multi-expert evaluations (including the aggregation of opinions and information from different sources);
- documenting HRA uncertainty assessment of quantitative results;
- compatibility of HRA with PSA; incorporation of only more important human failure events into the equipment-oriented logic structures (e.g. fault and/or event trees).

Some difficulties might be overcome if the HRA studies were to some extent computerized. These issues are discussed in detail in Sections 3 and 4.

2.6.2 Evaluation of advanced (qualitative) HRA methods

CES

The process of modelling of the plant is not exhaustively described. The simulations seem to be very complex and performed at a high level of abstraction. It is necessary to distinguish classes of NPP situations but the authors do not propose any concept. While CES may run on one or only a few incidents that are instances of the class, it would be necessary, in estimating the likelihood of a given contributor to intention failures, to aggregate the probabilities of all the ways these cognitively challenging situations could arise. In cases of the absence of empirical data the authors propose using multi-expert judgement techniques, e.g. SLIM-

MAUD procedure. No description was given of how to take advantage of such complex simulation or use CREATE methodology to support such judgements. It seems that the entire methodology is not easily applicable to support the engineering HRA/PSA studies.

Integrated modelling of operator cognitive behaviour

The modelling approach was not primarily developed for quantification of human reliability. The overall scope of the project was intended to support the analyses of man-machine interactions in accident conditions in order to: (1) Improve the system design; (2) Improve the operator training; and; (3) Improve the support given to the operator in cases of decisions.

The outcome of these simulations, as the authors propose, might be processed by an expert system that would enable one: to recognize novel information, to chain this information to already existing data bases and eventually to organize the rules to be supplied to the operator in a decisional support system (DSS).

COSIMO

The important peculiarity of the model is its ability to simulate the evolution of a human-machine interaction which is adaptive both to the internal state of the operator and to the history of the human-machine system. The complexity and flexibility of the simulation is assured by an AI software of the blackboard architecture. In the model the synergism of AI and psychology approaches was exploited. The model developed relies, however, on strong theoretical assumptions, and therefore makes more demands on validation. The validation phase thus represents the future necessary step for developing the model. An alternative would be to start from data collected in real situations and to formalize a model on this basis.

The methods described in Section 2.5 are not directly useful for quantification of HEPs. They were developed in principle for different purposes, e.g. to evaluate the quality of the man-machine interface and the influence of PSFs on human performance in the dynamic environment (stress, quality of operator training, quality of procedures, etc.). These methods seem to be useful for training purposes and designing the operator support system, e.g. through finding the potential for intention failures.

As was mentioned, there has lately been a call for developing a "third generation" of human reliability models /Do90/. Some researchers indicate that AI technology offers rich computer environments to better understand errors through the effort of modelling human behaviour, but proposed HRA methods should be based on a sound psychological theory foundation /Fu92/. Thus the methods assessed above (described in Item 2.5) can be also considered as important concepts and initial steps to develop advanced HRA engineering techniques for cases of event-driven situations in dynamic man-machine systems.

3 Concepts of the software system supporting human reliability evaluation and documentation

3.1 Methodological issues of HRA

3.1.1 Human interactions and assumptions concerning HRA/PSA

As described in Item 2.2 human actions/errors can be divided into three categories so as to relate to the phases of an accident (see Table 2.3).

Categories A to C of human interactions reflect the actions/errors made by personnel directly interfacing with plant controls and equipment. However these personnel are themselves dependent on other plant personnel and human influences which can significantly affect the resulting risks. The PSA encompasses these factors by making explicit or implicit assumptions about them and then modifying the assessment of the category A to C events to reflect the assumed effect of dependencies.

Maintenance activities (i.e. maintenance, test, repair, calibration and inspection) can lead to category A or B events. Control actions/errors during normal operations can lead either to events of category B or A, e.g. failing to realign systems/equipment after testing; omitting to restore protection system vetoes or permissives. Events of category C can occur either in the control room or locally at the plant /IA92b/. The two categories A and B are behaviourally identical, differing only in whether the results of the failure immediately upset plant phenomena or not /Do88/.

Actions of types C1, C2 and C3 (Section 2.2, see Table 2.4) are all involved in operator response to an accident once it has been initiated. Often, some diagnosis is required and time is a limiting factor. The general approach for dealing with type C1 and C3 interactions is the same /IA92b/.

Potential interactions C2 should be identified jointly by human reliability analysts and systems analysts. A first type of such interactions occurs when the operator's mental image of the plant differs from the actual state leading the operator to perform the right actions for the wrong event. Another type occurs when the operator correctly diagnoses the event, but chooses a non-optimal strategy for dealing with it. Only a few PSA studies have attempted to include this type of interaction, and only to a limited degree /IA92b/; see also Section 3.4.

Events associated with actions/errors of categories A-C are to be incorporated into fault trees and event trees. An outline of the relevant procedure is described in Item 4.1. Actions of category C3 can be incorporated as recovery factors on the frequency of the accident sequence cut sets /IA92b/.

Examples of assumptions normally explicitly recorded and those that are usually implicit in a PSA are as follows /IA92b/:

(a) EXPLICIT

- level of control room staff experience and general training,
- type inspection and QA (Quality Assurance) controls applied to maintenance,
- comprehensiveness and accuracy of procedures (operations, maintenance and testing),
- quality of man-machine interfaces,
- quality of communication links between personnel "on the plant", e.g. between control room and those local to the plant.

(a) IMPLICIT

- ability to update procedures with plant modifications during plant lifetime,
- sensitivity/robustness of organization to changes of plant personnel (particularly key management positions),
- adherence of staff to procedures, technical specifications, QA procedures,
- ability and effectiveness of the management and organization to resolve safety and economic conflicts acceptably,
- ability to communicate important information through the organization to reach the appropriate people.

3.1.2 Types of human behaviour and errors

Rasmussen's distinction of three categories of human behaviour is useful in evaluating of human reliability. A conceptual framework illustrated in Fig. 3.1 assumes three levels of human behaviour: (I) Skill-based; (II) Rule-based; and (III) Knowledge-based. These human behaviour types are described in Section 2.2. Skill-based tasks require little or no decision making and hence the task goes from initial stimuli, i.e. "Activation" to "Execution". Rule-based tasks require some decision making to proceed from the initial stimuli to the "Integration" phase where information is processed, then an appropriate procedure or rule selected. Finally, the selected procedure or rule is executed.

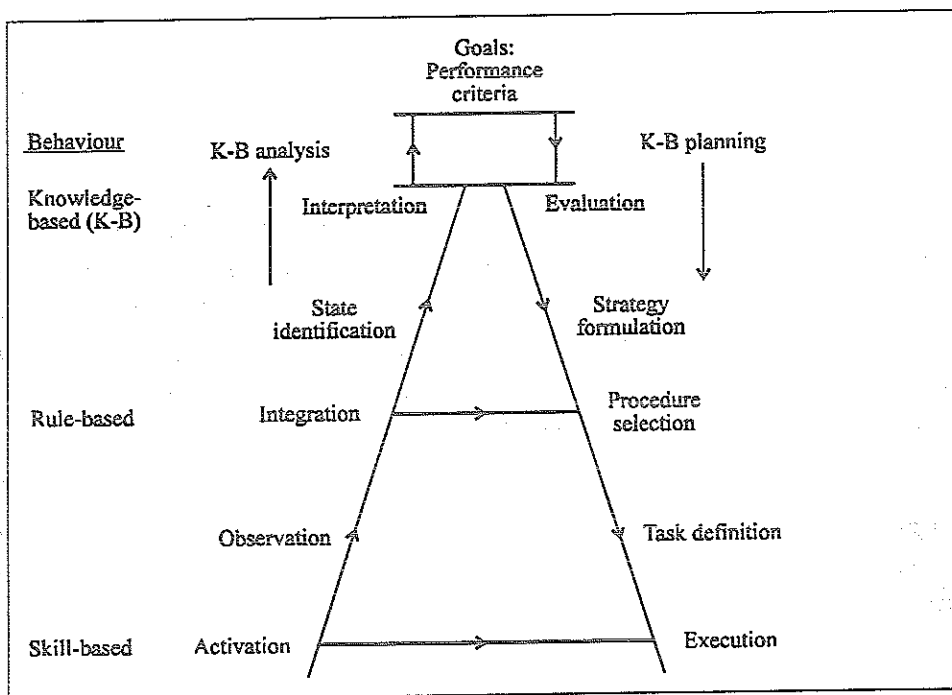


Fig. 3.1: Rasmussen's conceptual framework of human behaviour and decision making

Knowledge-based behaviour involves cognitive activities such as planning and a comparatively high degree of decision making. The relevant process of human behaviour includes the interpretation and evaluation of information, planning of actions and the execution of a plan.

The plan can consist partly of some appropriate procedures or rules. Knowledge-based behaviour is often necessitated by uncertainty or novelty in the circumstances at hand.

Reason's /Re87/ cognitive framework for locating common human error forms GEMS (Generic Error-Modelling System) corresponds to Rasmussen's framework. The main features of GEMS are summarized in Fig 3.2. Its operations divide conveniently into two areas: those which precede the detection of the problem (skill-based level); and those which follow it (the rule- and knowledge-based levels). Errors (slips) occurring prior to problem detection are primarily associated with monitoring failures, while those which appear in the subsequent phases (mistakes) are subsumed under the general heading of problem-solving failures.

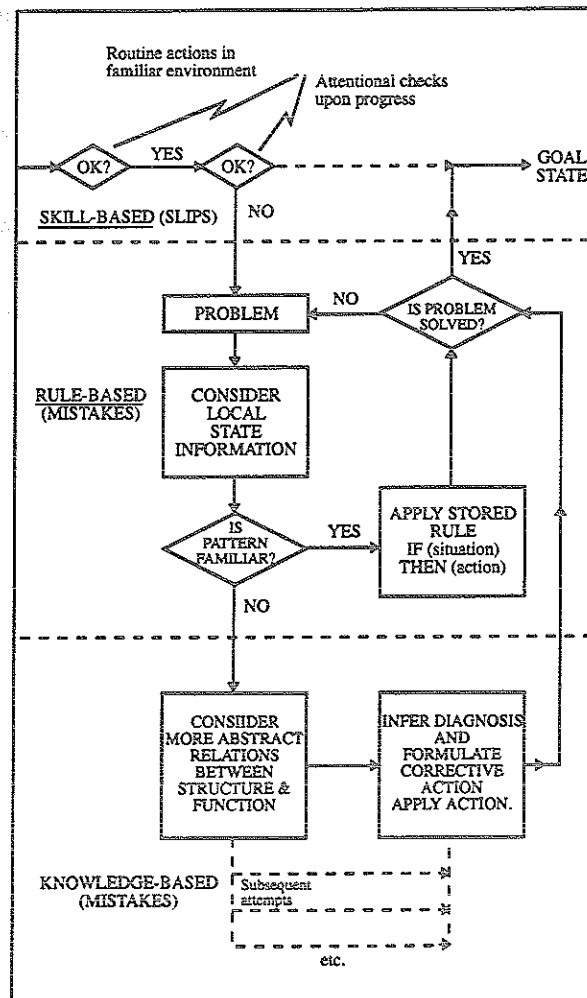


Fig. 3.2: Dynamics of Generic Error-Modelling Systems (GEMS) /Re87/

Monitoring failures and other routine actions, in their simplest form (i.e., well practised tasks carried out in familiar surroundings by skilled operators) can be regarded as comprising segments of preprogrammed behavioural sequences interspersed with attentional checks upon progress, carried out either consciously or preconsciously. These checks involve bringing the higher levels of the cognitive system momentarily into the control loop and seeking to establish: (a) whether the actions are running according to plan; and (b) at the more complex level, whether the plan is still adequate to achieve the goal. Studies of naturally occurring

action slips clearly indicate the dependency of slips upon failures of attentional checking. These failures are basically of two kinds: the large proportion are associated with inattention (omitting to make a check), but a significant number of slips are also related to overattention (making an additional check at an inappropriate moment in the sequence /Re87/).

Problem-solving tasks usually require new plans and strategies to be derived from the 'first principle'. The problem-solving elements of GEMS are based upon a recurrent theme in the psychological literature, namely that humans, if given a choice, would prefer to act as context-specific pattern recognizers rather attempting to calculate or optimize. Thus the sequences of operations within GEMS indicates that human problem solvers always confront an unplanned for change situation by deciding whether or not they have encountered this particular pattern of local signs and symptoms before. If pattern is recognized, they then determine whether or not they have some corrective rule readily available. Only when this low-effort pattern-matching and rule applying procedure fails will they move to the more laborious mode of making diagnostic inferences from knowledge-based mental models of the system in question, and, from these, go on to formulate and try out remedial possibilities /Re87/.

There are some 'switching rules' governing the shifts of control between the skill-based, rule-based, and knowledge-based levels represented in Fig. 3.2. With more difficult problems, the cycle skill-based $\leftrightarrow$ rule-based (scanning local signs and symptoms, rule matching, and evaluating solution effectiveness) may be repeated several times. Also for more complex problems the cycling between the knowledge-based and rule-based levels can be repeated several times as various theories are explored. If the local signs and symptoms are changed as a consequence of the previous activity, then in some cases the focus of control will shift between all the performance levels /Re87/.

Reason assumes in GEMS that certain error-shaping factors are intrinsic to the way the cognitive system functions, and that human beings are prone to make certain kinds of error irrespective of the task or the situation. He distinguishes intrinsic and extrinsic factors in shaping specific error forms. His tentative guesses are given in Table 3.1. Intrinsic factors include cognitive biases, attentional limitations, and limitations to human rationality. Extrinsic factors include task characteristics, context effects, and factors from the environment.

	INTRINSIC	EXTRINSIC
SKILL-BASED SLIPS	High	Moderate
RULE-BASED SLIPS	Moderate	High
KNOWLEDGE-BASED MISTAKES	Low	Very high

Tab. 3.1: Contributions of intrinsic and extrinsic factors in shaping specific error forms /Re87/

Thus GEMS confirms the difficulty known from HRA practice that the distinction between skill-based actions and rule-based actions is often more or less arbitrary. A similar difficulty is often to be faced with the distinction between rule-based and knowledge-based behaviour. Moreover the cognitive level of operator behaviour depends on his training and experience (novices perform almost all interactions on the knowledge-based level). Therefore in cases of

doubt concerning the cognitive level required for a given interaction, the higher level should be assumed (with higher HEP).

At the knowledge-based level mistakes can take a wide variety of forms, none of which is necessarily predictable on the basis of the individual's past experience. To achieve more precise qualitative predictions at this level, we need a better understanding of the complex interaction between task variables and problem solving strategies, and of the way in which the latter change as a consequence of continued efforts at finding a solution /Re87, Re90a/.

Mistakes are associated with rule- and knowledge-based actions. They are a more serious error mechanism as they lead to e.g. incorrect understanding of a situation and the selection of an inappropriate plan of resulting actions (i.e. sequences of actions). Mistakes can occur due to inappropriate selection of familiar procedures or rules based on only some of the available information, in order to reduce the amount of decision making required /IA92b/.

Slips usually occur in a less time-constrained environment, it is, therefore, reasonable to believe that their associated error probabilities are less time-dependent, more structural, and arise as faults in an intended protocol or procedure. On the other hand, mistakes occur more often in a time-constrained environment and are more time-dependent (it takes "process" time to think up an appropriate response in a novel situation). Mistakes are less structural, and often have little to do with protocol or procedure /Do88/.

3.1.3 Recovery potential

For PSA purposes it is important to be aware of the error recovery potential for different types of error mechanisms. Slips and lapses can generally be recovered fairly quickly providing there are appropriate feedback mechanisms and the plant behaviour is reversible. On the other hand, mistakes are less easily recovered in the short term. "Mindset" problems can occur and operators can persist in attempting to implement an inappropriate plan even when faced with a lot of contradictory information. Recovery routes generally have to be very positive and powerful to be reliable /IA92b/.

If an error of category A (Table 2.3) is made, the significance and impact on the system unavailability or failure rate depends essentially on two aspects: (1) how quickly the error is discovered and corrective actions taken, and (2) what has been done to minimize the possibility that the same error is repeated in other redundant components or in other systems (possible common cause failures). Factors that influence one or both of these features are called recovery factors. Some factors mainly affect the error discovery rate, other mainly the likelihood of repeated (dependent) errors. One should identify which of them have a potential to influence actions/errors that are considered in the study. Factors affecting errors of category A and B (Table 2.3) are as follows /IA92b/:

Recovery factors affecting mainly the error discovery rate:

- RA - an automated system that realigns a train or component for performing a safety function even if valves were left in the wrong position;
- RB - a clear alarm annunciator or a significant change in a parameter that will readily alert operators about a wrong status;
- RC - an independent test after calibration, limit setting or maintenance that discovers an error (and does not by itself require any disconnections or re-alignments);

RD - a requirement and practice that another person inspects the component status and/or electrical connections after the original maintenance or calibration, with a check list or sign-off obligation, or that the original performer makes a separate inspection with similar obligations;

RE - a daily or periodic inspection is carried out that would discover the error, and there is a check list or sign-off obligation specific for the task error.

Recovery factors affecting mainly the likelihood of repeated (dependent) errors:

DA - a strong administrative control of keys and locks on redundant components and scheduling and authorization of work on the task activities to assure that one person or crew does not work on more than one redundancy train at a time;

DB - a requirement and practice to replace inspection tools, chemicals and calibration instruments between tasks on different redundant trains;

DC - the task for different redundancies can be carried out in different rooms (to help minimize the temptation to collect work orders and work on multiple redundancies consecutively);

DD - a practice that assures sufficient time between actions for different redundancies, if one person or crew is allowed to work on all redundancies;

DE - an obligation to check or write something down on paper after each redundancy, before proceeding to the next.

Remark: the analyst should identify other plant-specific recovery factors as well.

3.2 Classification of human actions/errors and quantification techniques

3.2.1 The HRA procedure

PSA makes certain requirements on HRA. To satisfy these requirements, the HRA must provide assurance that /Do88/: (1) all risk-significant human events are identified; (2) these events can be incorporated into the PSA models; and (3) these events are quantifiable with probabilities. In Item 2.4 two frameworks for incorporation of HRA into PSA are described, namely: SHARP /Ha84/ and a recent one proposed by a group of HRA/PSA experts asked by the IAEA /IA92b/. A simplified scheme of some important steps and tasks of the HRA process is illustrated in Fig. 3.3.

The principal strategy in reliability analysis is to decompose failure events into meaningful subevents for which there are available data to provide estimates of some reliability characteristics. The purpose of classifying failure events is to uncover these reliability-related subevents and to form the basis for collecting or understanding data concerning these identified subevents. This classification process should be based on physical principles, e.g. failure analysis, or at least observation. The more realistic the foundation for the classification scheme, the greater the predictive power and credibility that the classification can claim. HRA also adopts this classification strategy /Do88/.

A properly engineered PSA approach is characterized by being integrated and economical. Correspondingly, the HRA tasks are partitioned so as to match the PSA needs. This results in three groups of tasks: (1) integration tasks; (2) tasks that support the initial quantification(s); and (3) tasks that provide the detailed event analyses at the end of the PSA /Do88/. One of the most important steps of the analysis is the event characterization and selecting of the HRA technique, which is considered in the next section. Other issues concerning the incorporation of human-induced events into PSA logic are discussed in Chapter 4.

3.2.2 Human actions/errors and selection of quantification techniques

Section 3.2.2 includes an extended description of an approach outlined in /Ko92/.

A classification scheme for human failure events suggested in ref. /Do88/ and presented as the "decision" tree leads to several categories of human failure events. This taxonomy can be easily incorporated into a PSA and is consistent with others proposed, such as that of Embrey /Em86/.

The approach proposed should be in agreement with methodological aspects of HRA discussed in Section 3.1. Two classification proposals are shown in Fig. 3.4 and 3.5. It was assumed that the degree of preparation expended on an activity can be the categorizing factor. Human interactions can be divided into pre-incident (plan-driven activities) or event-driven activities. These correspond to categories A and C (see Tab. 2.3) actions/errors and are respectively represented in Fig. 3.4 and 3.5. These selection trees are extensions of a concept of the techniques selection described in /Do88/.

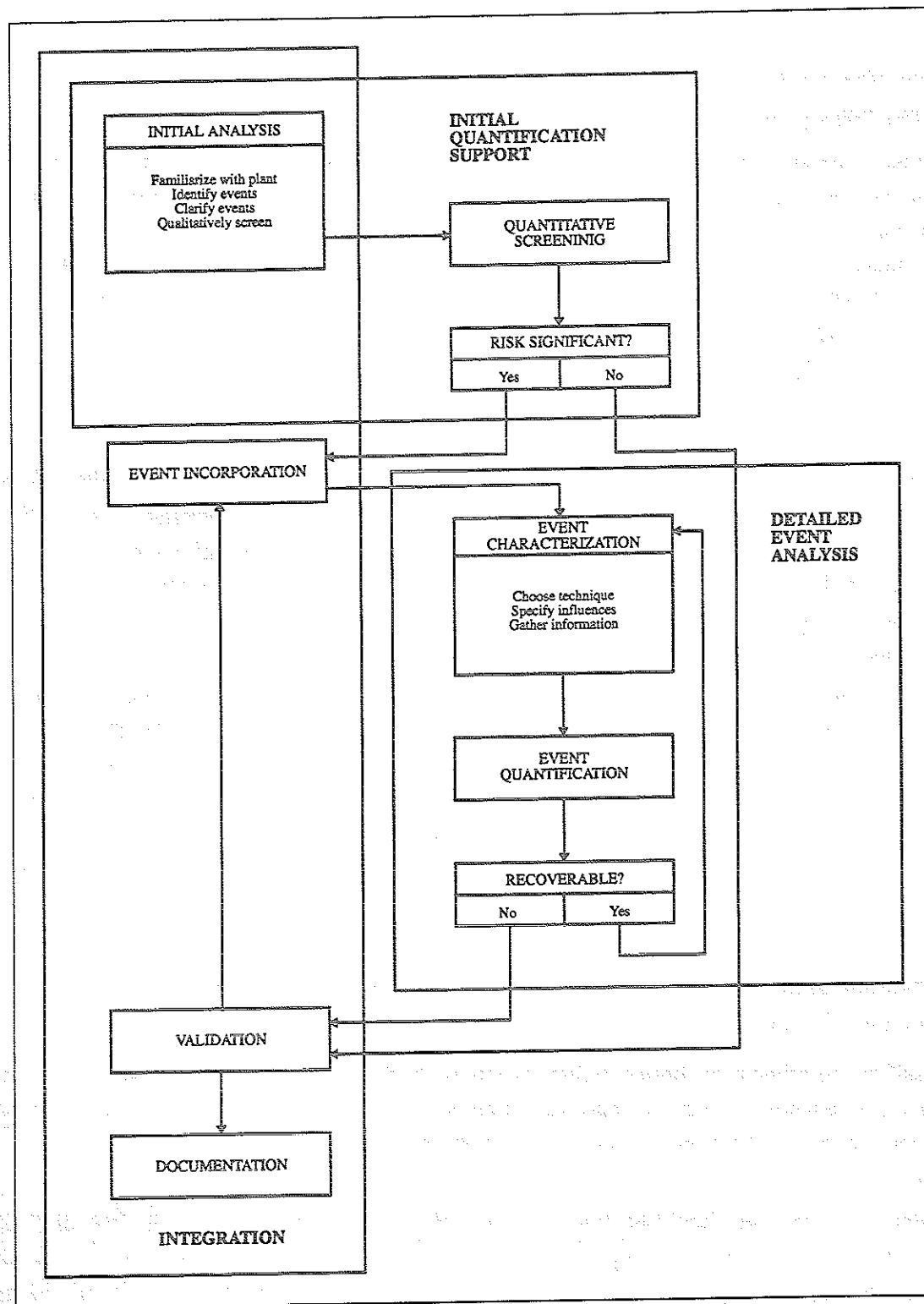


Fig. 3.3: The HRA process /Do88/

Plan-driven activity is implemented according to a specific procedure for an established task. Proceduralized actions in routine task activities are prescribed by supervision or work order and there is little opportunity for mistakes; thus, slips dominate (Fig.3.4).

Action	Error type	Effect	Category	PSA/HRA effort			Quantification techniques
				I	II	III	
Specificable task	Slip	Omission	P1				SLIM, PC, APJ THERP ASEP-HRAP
		Commission	P2				SLIM, PC, APJ THERP ASEP-HRAP
			P3				APJ, (PC) data, if at all not modelled
	Mistake	Omission	PR1				APJ, PC THERP ASEP-HRAP
		Commission	PR2				APJ, PC THERP ASEP-HRAP
			PR3				APJ not modelled not modelled
Recovery							

Fig. 3.4: Classification proposal for human pre-incident actions/errors and quantification technique

Event-driven activity depends on an off-normal situation at hand and may also be anticipated by procedure, e.g. the so-called symptom-based emergency procedures. Failure in this activity can be referred to as response failures, in that a human response was anticipated and designed for. Recovery is more ad hoc and less directly supported by procedures. Almost all of the resulting event types could include either slips or mistakes as error types /Do88/ (Fig. 3.5).

Mistakes are not usually modelled in normal task activities (Fig. 3.4, category P3). The driving assumption is that the plan is specified by plant management and directed by technical specifications and the decision elements of the task are, thus, highly reliable. However, the Chernobyl accident in 1986 indicates that there may be circumstances under which this assumption might need to be reevaluated /Do88/.

Commission failures that are due to mistakes are not always modelled (Fig. 3.5, categories E4 and ER4). This is because it is assumed that other failure modes dominate their occurrence probabilities and the presence of extensive plant instrumentation makes their effects detectable and gives rise to the possibility of mitigating the consequences of these failures. However, this assumption can be questioned in the case of events degrading the state of the plant and depending significantly on time /Do88/.

During off-normal conditions, awareness is heightened, more than a single person is around, but time is usually an important factor. Thus, slips will not often go unrecognized or uncorrected, and diminish in significance. Mistakes, however, can occur due to confusion, uncer-

tainty, conflict, or lack of knowledge, and the effects of mistakes may be universal and persisting /Do88/. Thus, in such a situation mistakes will be more likely than slips. In other words, during interactions of category A, B and C1 (see Tab2.3 and 2.4) slips are more likely to be made than mistakes. On the other hand, actions of category C2 and C3 are rather dominated by mistakes than slips.

Action	Error type	Effect	Category	PSA/HRA effort			Quantification Techniques
				I	II	III	
Response	Slip	Omission	E1				SLIM, PC, APJ
							THERP
		Commission	E2				ASEP-HRAP
							SLIM, PC, APJ
	Mistake	Omission	E3				THERP
							ASEP-HRAP, INTENT
		Commission	E4				SLIM, PC, APJ
							TRC, HCR
Recovery	Slip	Omission	ER1				TRC (response)
							CM+APJ
		Commission	ER2				CM (MS)
							ASEP-HRAP
	Mistake	Omission	ER3				APJ, PC
							THERP
		Commission	ER4				ASEP-HRAP
							APJ, PC

Fig 3.5: Classification proposal for human event-driven action/errors and quantification techniques

The current state of the art of human reliability analysis indicates, as has been described above and in Chapter 2, that different techniques need to be applied to cover the probabilistic evaluation of various cases of human failure events. In addition during the selection of HRA techniques other factor must be taken into account (Table 2.1):

- accuracy, consistency and validity of techniques,
- usefulness, maturity and acceptability of techniques,
- effective use of resources (equipment and personnel resource requirements, data requirements, training requirements, effort required, etc.).

3.2.3 HRA/PSA effort and selection of quantification techniques

Effort devoted to HRA, associated with the scope of analyses and the number of experts involved, should correspond with planned effort to perform PSA. For supporting the HRA techniques classification and selection it is proposed to distinguish three scopes of PSA/HRA:

- (I) Small effort PSA/HRA;
- (II) Medium effort PSA/HRA;
- (III) High effort PSA/HRA.

Features of PSA for these three efforts are described in Item 4.2. The scope of HRA should be balanced with PSA effort. Therefore, in Fig. 3.4 and 3.5 different HRA techniques are proposed at three levels. For the first level such simplified techniques were selected for quantifying relevant cases of actions/errors as ASEP-HRAP, INTENT, TRC (generic). The techniques selected for the second level require higher resources: THERP, HCR, specific TRC, but in principle HRA can be performed by a single expert to prepare e.g. initial results to be considered in HRA performed with the contribution of several experts (level 3).

To limit subjectivity of estimates given by experts some modification of basic methods can be proposed, e.g. in modelling of diagnosis error the probabilities of the confusion matrix can be evaluated with regard to similarity measures of the plant responses obtained from deterministic modelling for different initiating events. Such an approach, called Confusion Matrix with Modelling Support CM(MS), is outlined in ref. /Ko92/.

At the level three the multi-expert techniques are foreseen to deal with most important and difficult cases which require combining of expert opinions. Such cases will be identified in HRA/PSA performed at lower levels.

Thus, the evaluation of human error probabilities can be carried out in the following basic steps:

- (1) Assigning the category of human failure event based on the classification proposal presented in Fig. 3.4 and 3.5 taking into account attributes foreseen in the classification system;
- (2) Selection of the HRA technique (or techniques) corresponding to assumed HRA/PSA effort (human reliability data obtained from simulator experiments can be included through THERP or TRC techniques); some techniques require supporting simulation of the plant response to assess a time window for the operators' reaction (THERP, HCR) or some similarity measures (CM with MS);
- (3) Estimation of the human error probability with regard to the situation attributes (including PSFs) relevant to the technique selected (for more complex cases of diagnosis made by operators, using the OAT technique can be useful);
- (4) Considering potential for recovery of human errors.

The third step can be quite complex including, e.g. in the case of the THERP technique, sophisticated searching in many tables of data. This searching and calculation of probabilities according to mathematical formulas can be very complex for some cases and require a high

level of expertise. An expert system supporting and documenting HRA using this technique is described in Section 3.5.

3.3 Quantitative analysis of human reliability

3.3.1 General remarks

Some issues to be dealt with in the quantification of human failure events are as follows /Do88/:

- (1) Human actions/errors may include human contributors to initiators, if initiators are modelled rather than quantified directly from data. Slips may or may not be included in rule-based events. Recovery events are usually attributed only to mistakes. Event specifics are identified by informal or formal interviews of operators and plant staff or observations in simulator exercises, whenever possible.
- (2) The events are represented and incorporated into the PSA structures (fault trees or event trees) either directly in a logic structure or indirectly, using first Human Reliability Analysis Event Trees (HRAETs) or Operator Action Event Trees (OAETs) to evaluate the success/failure probability of events of interest. The incorporation of modelled human-induced events into fault/event trees is discussed in Item 4.1.
- (3) Usually slips are modelled using a simplification of THERP approach; mistakes can be modelled using a combination of TRCs and SLIM.
- (4) Probabilities and their distributions are obtained according to the methods of the chosen technique.

An important step in HRA procedure is quantitative screening. Some probabilities for slips, taken from THERP, range one or two orders of magnitude around the value of 0.001. This value can be used as a basic screening value for latent or human-induced initiator events. THERP also provides a value for the conditional probability for a second error given the first. Assuming moderate dependency, this value is about 0.1. The product of these two values can be used as a screening value for the human failure resulting in the failure of multiple (particularly redundant) components /Do88/.

3.3.2 Quantifying slips

Slips can occur in routine process plant activities, for example, when a maintenance person is returning a piece of equipment to service and fails to leave it in the proper state. Both auxiliary feedwater valves at TMI were apparently left closed by the maintenance team, which rendered a vital safety system unavailable.

In the case of a slip, there is usually no unanticipated event to respond to and planning is specified by procedure, policy, or work order. It is fair to assume, then, that the main cognitive element is that of control, i.e. carrying out all steps of the plan. There are many potential mechanisms of slips that THERP assumes because each step of a particular procedure is a candidate for a slip (i.e. an omission or commission). However, the likely candidates for slips in a control room seem to be limited to /Do88/:

- (1) Stereotype capture, i.e., a familiar, more frequently performed task uses some of the same steps as the intended task, and in the first case of deviance, the operator slips into the old task rather than continuing what is called for in the required task;
- (2) Spatial reversals, e.g., two controls are adjacent and the wrong one is manipulated because of a lapse in attention (or deviation from stereotype);
- (3) Time reversals, where the sequencing of manipulations is critical and the order is incorrect, again due to a lapse.

A more detailed list of slip subcategories is given in /Re90a/; see also Table 3.2 in this report.

Assuming that the mechanisms of slips are limited to the above, only one instance of a slip per task should be modelled, which is the difference from THERP. A review of the estimates for task element probabilities in Chapter 20 of NUREG/CR-1278 /Sw83/ shows a range from 0.0003 to 0.003 for most of the types of situations relevant to process activities. Thus, a simple strategy is adopted of using the geometric mean, 0.001 of the bounds, as the basic probability of any slip; 0.0001, or an order of magnitude lower, for a slip that leaves redundant equipment inoperable. This is associated with the screening mentioned previously/Do88/.

THERP, then, allows the basic estimate to be adjusted by PSFs based on situational, task, or behavioural influences. For example in maintenance, there may be a check off policy and a PSF should give credit for this. The possibility of human engineering deficiencies can also be accounted for by a PSF, e.g., both valve operators for two trains are adjacent or reversed or otherwise could induce a reversal slip /Do88/.

Doughtery /Do88/ suggests (based on Table 20-22 in /Sw 83/) that the first factor should be 0.1 if the checking is independent of the manipulation. If a human engineering deficiency exists, then the second factor could be greater than 1.0, say, 5.0. The net result for the slip, then would be: event probability = basic probability x factor1 x factor2 = $0.001 \times 0.1 \times 5 = 0.0005$ /Do88/. Different factors might be proposed to account for various cases of recovery factors distinguished in /IA92b/ specified above in Item 3.1.3.

This simplified THERP approach may be used for slips whether they are latent, response, or recovery events. A more detailed approach was developed for the Oconee PRA. The treating of human-induced dependent failure requires especially careful consideration /Do88/. Some additional explanation concerning quantifying of slips is presented in Item 3.5 (description of an expert system for the application of THERP).

3.3.3 Quantifying mistakes

A mistake results from one of the following failure modes: misdiagnosis, faulty decision, or faulty planning. These failures modes seem to matter most when the process has deviated from normal and desired conditions, i.e., in an abnormal or emergency event. The diagnosis and decision making of a kind needed in off-normal response is influenced by available time, perceived time (or the pace or urgency of events), uncertainty, complexity, and goal conflict. These latter three factors are the primary sources of the phenomenon called burden /Do88/.

The individuals involved in diagnosis and decision making often exhibit hesitancy. Hesitancy can be due to uncertainty in the conditions present, thus inhibiting or failing to allow for proper diagnosis, or can be due to uncertainty as to which goals to pursue when one or more appear to conflict. The conflict can be exhibited even when relatively clear procedures exist for these conditions. For this reason, diagnosis and decision making seems to be the critical and dominant risk-significant human behaviour in an abnormal or emergency situation /Sw83, Do88/.

Because errors in diagnosis and decision making can have such serious consequences, nuclear power plants have attempted to eliminate or, at least minimize, the role of diagnosis and decision making by introducing so-called symptom-oriented emergency procedures. These procedures contain "rules" of the form /Do88/:

IF symptoms $S_1, S_2, \dots, S_n$ exist
 THEN execute actions $A_1, A_2, \dots, A_m$
 AND do so without hesitation.

In most cases, the symptom set is kept small and each symptom is clearly indicated in the control room. However, there can be cases when symptoms are similar and due to confusion the operator can activate the incorrect procedure /IA92b, Wa88a, Ko92/.

These aspects of cognitive interactions and possible errors motivate the development of a model of process control mistakes that has the following characteristics /Do88/:

- (1) Time is an independent variable; the probability of successful diagnosis or decision making (D&D) increases with available time;
- (2) Crew effect is aggregated, i.e., the failure probability estimates are for an anticipated crew structure, such as defined in NUREG/CR-1278, Table 18-2, not for any individual;
- (3) Conflict, complexity, or other sources of burden are prime influences on the reliability of a D&D performance. This influence is explicitly factored. The presence of rules is also explicitly factored;
- (4) Other influences, such as the quality of procedures, the adequacy of the instrumentation and controls, and the adequacy of training, can be systematically factored into quantification using SLIM, or some other subjective judgement approach;
- (5) Slips are not usually dominant when D&D is necessary, or when time is forgiving, or when multiple crew members are present to notice and correct slips. In cases when slips are judged to be significant, the THERP approach can be applied to misaction and the THERP annunciator model (Chapter 11 of NUREG/CR-1278) can be applied to misdetections, if postulated as important;
- (6) Early, incorrect action based on inadequate D&D (sometimes also referred to as a commission error) is assumed to be correctable, in principle. The likelihood of the error is subsumed in the TRC value. Its effect are not modelled unless they would change the course of the sequence radically.

Different modelling methods can be considered depending on the case of mistake analysed (Fig. 3.5), i.e. error of commission or error of omission: TRCs (generic, or plant specific in the case of data available from the simulator), CM (also CM with modelling support: CM-MS /Ko92/), HCR, and SLIM as an interpolation tool.

In reference /Do88/ two pairs of curves in the family TRCs were distinguished (Fig. 2.2). One pair is applied to response or rule-based mistakes, i.e., misdiagnosis or faulty planning or decision making when guided by the kind of rules (one curve when hesitancy is not important influence; the other curve when hesitancy is dominant). The second pair of curves is applied when on-the-spot, general diagnosis must be used to decide on a course of action in the absence of rules. This will most often be the case for recovery events. Again two curves are used to recognize hesitancy or its insignificance. The recovery curve is a lognormal fit of the TRC (which was used in THERP to apply to the diagnosis of choosing an appropriate procedure). THERP did not include a curve for rule-based responses.

A generic calculation of the probability of an unrecovered commission error can be represented by /Do88/:

$$P = P1 \times P2 \times P3$$

where: $P1$ is the probability of an extended and significant commission error; $P2$ is the probability that Emergency Response Guideline do not cover the condition resulting from the

error; and P_3 is the probability of recovery by the SRO (Senior Reactor Operator) or other personnel. It is suggested that the value of P is very low and is bounded by other failure modes /Do88/.

The next problem with the TRC system is that lognormal CCDFs do not terminate at some low probability. However, the probability of unrecovered slips, even with multiple crew present, can serve as a lower threshold. It has become standard practice to truncate failure probabilities generated by TRCs at 0.0001 or 0.00001, or truncate the available time at about an hour, as an intermediate screening process /Do88/.

Other factors influencing the possibility of human error such as the adequacy of the control room's instruments and controls, the applicability of procedures, the communication of the crew, etc., can be considered using the SLIM technique as an interpolation tool.

Mistake with an effect of commission can be modelled using the CM technique (Item 2.3). Modifications of this technique are proposed to account for plant-specific rules concerning the activation of operational procedures /Wa88a/ or similarity measures of the plant response for initiating events considered (CM with modelling support: CM-MS) /Ko92/. The purpose of these modified approaches is to support experts in quantifying of errors, to limit subjectivity of estimates.

HCR technique is often used for modelling mistakes with an effect of omission (see remarks on application of this technique in Item 2.3). The last two methods discussed must be, however, combined with other techniques to account for other aspects of human reliability modelling. Also in these cases the SLIM technique can be used as the interpolation tool. However, the contribution of several experts is required. There is an obvious need to support the human reliability analysers with a computer program to calculate HEPs and to document HRA (see description of ORCA in Item 2.3).

One of the most challenging problems in the probabilistic modelling of human behaviour is cases of conflicting goals, i.e. which goal to pursue when one or more appear to conflict in off-normal process plant events. Unfortunately for solving such problems no suitable methods have been developed within the decision making theory /Do88/. A similar problem to be faced is the modelling of intention failures. The methodology proposed by Woods /Wo87, Wo88/ is not verified and difficult for engineering applications. More practical from the HRA/PSA perspective seems to be the INTENT method that, however, should also be verified /Ge92/. A recently developed method is described in Section 3.4 of this report.

3.4 Proposed method for the quantification of decision-based errors in HRA

3.4.1 State of research

This section contains a brief description of a model developed in /Re92b/ for the quantification of decision-based errors of operators in an NPP.

Current HRA research emphasizes intentional errors in decision-making. Such errors will result in extraneous acts which reduce system safety directly or indirectly (with an increased probability). Here, a method is proposed which allows the HRA quantification of decision-based errors due to risk taking behaviour in situations with conflicting goals; e.g.:

Nuclear power plant Davis-Besse, June 9, 1985, about 10 min after loss of feedwater. The procedure requires the initiation of feed & bleed cooling (given this cooling mode, coolant would be released inside the containment building, and cleanup requires the plant to be shut down for days). Although the shift supervisor had diagnosed the loss of feedwater event correctly, he did not initiate the feed & bleed cooling. However, this decision resulted in a positive outcome: about 4 min later feedwater was restored /NU85/.

According to /Re90a/, such behaviour can be classified as "violation" (see Tab. 3.2).

UNSAFE ACTS			
UNINTENDED ACTION		INTENDED ACTION	
SLIP ↓	LAPSE ↓	MISTAKE ↓	VIOLATION ↓
Attentional failures Intrusion Omission Reversal Misordering Mistiming	Memory failures Omitting planned items Place-losing Forgetting intentions	Rule-based mistakes: Misapplication of good rule Application of bad rule Knowledge-based mistakes: Many variable forms	Routine violations Exceptional violations Acts of sabotage

Tab. 3.2: A summary of the psychological varieties of unsafe acts: /Re90a, p.207/

Risk-taking behaviour is related to decision-making in hazardous situation. Little effort has been made to operationalize and observe the individual regulation process in such situations /Ho88, p.101/.

Current HRA methods are not sufficient for the situation-specific quantification of decision-based errors due to risk-taking behaviour. The most commonly accepted THERP method gives no specific HEPs related to "postdiagnosis decision making" /Sw83, p.12-8/, although, the problem of conflicting goals is recognized /Sw83, p.12-9/.

Decision based errors are the subject of the INTENT method /Ge92/. Interval estimations from experts are given related to the probabilities of 20 so-called basic error types. Eight types are related to risk-taking behaviour; e.g.: "Violate procedure and devise own formulae". To estimate the probability of a selected situation-specific error type, INTENT proposes a SLIM-

based technique for interpolation within the probability interval of a selected error type. It requires ratings to 11 PSFs (e.g. "Human-Machine Interface"). However, only a few PSFs are related to risk-taking behaviour. Essential aspects are not included; e.g., the possible outcome of risk taking or safe behaviour (see /Ed68/).

Several HRA methods have recognized the importance of PSFs related to risk taking behaviour in decision making. E.g., referring to the Davis-Besse event /Wo87a, p.80/, the CES method /Wo90b/ considers the requirement of an action departing from the usual as a "complicating factor" (see Tab. 3.3). However, this factor is not a component of the current CES version. The OAT method recommends increasing an HEP by a factor of 3, if there is reason for reluctance to perform the respective action /Ha82, p.26/.

There is no consensus on the quantification of such reluctances (related to risk-taking behaviour) in HRA. E.g., one case study of THERP /Sw83, ch.21/ assesses a basic probability (median) of 0.001 for non-initiation of feed & bleed procedure, given a correct diagnosis: "If the problem has been diagnosed correctly, it is very unlikely that the operator would fail to realise that help was needed". The quantification of recovery factors results in an overall HEP of 0.0000036 for this subtask (feed & bleed initiation) /Sw83, p.21-9/. Another study /Wa87, p.112/ assesses an overall HEP of 0.01.

In view of this state of research, here a method is presented which allows for situation-specific estimation of decision-based errors (due to risk-taking behaviour).

1. Human execution errors.
2. Additional machine failures (e.g., valves that stick open, systems that fail to work as demanded)
3. Missing information (e.g., sensor failures).
4. Multiple major faults (e.g., tube rupture with an unisolatable steam release from the faulted steam generator).
5. Situations which remove or obscure the unusual evidence or critical evidence (e.g., a loss of leading indicator incident such as a loss of offsite power prior to a steam generator tube rupture).
6. Complex situation where different parts of the situation suggest responses which conflict with each other (e.g., the Ginna incident).
7. Situations that require actions which depart from the usual (e.g., total loss of feedwater).

Tab. 3.3: Complicating factors - some variations or difficulties that go beyond the standard method for handling the situation: /Wo90, p.44/

3.4.2 Description of the SLIM-based model

"Blame and guilt tend to be associated with errors, and therefore many erroneous actions are not reported because of the likelihood of punitive actions against the operator" /Em84, vol.I, p.2/.

Facing this problem, here a model is selected which allows the quantification of risk-taking behaviour based on the sparse data currently available. It is assumed that the probability p of a decision based error can be calculated as follows:

$$p = p_o \prod_{i=1}^n PSF_i = p_o \prod_{i=1}^n K_i^{x_i} \quad (1)$$

where

$$0 \leq x_i \leq 1,$$

$$K_i > 1,$$

n = number of relevant PSFs,

p_o = minimum basic value of p ,

K_i = maximum value of PSF_i ,

x_i = rating ("degree of presence" related to error likelihood) of PSF_i .

This is a geometric quantification of a PSF effect to an HEP:

$$PSF_i = K_i^{x_i} \quad (2)$$

The range of x_i goes from 0 ("not present", i.e., best condition related to reliability) to 1 ("totally present", i.e., worst condition related to reliability). An x_i value of 0 results in a PSF_i of 1 (i.e., no HEP increase); an x_i value of 1 results in a PSF_i of K_i (i.e., maximum HEP increase).

Eq. (1) is equivalent to the SLIM model /Em84; Em83/. In /Em83, p.15/ an error probability p is modelled as follows:

$$\log p = C + D \cdot SLI \quad (3)$$

$$SLI = \sum_{i=1}^n (y_i w_i) \quad (4)$$

$$\sum_{i=1}^n w_i = 1$$

where

$$0 \leq y_i \leq 1,$$

$$0 \leq w_i \leq 1,$$

$$0 \leq SLI \leq 1,$$

n = number of relevant PSFs,

w_i = weight (relative importance) of PSF_i ,

y_i = rating ("degree of presence" related to success likelihood) of PSF_i ,

SLI = quantity which represents the overall belief of the judge(s) regarding the positive or negative effects of the PSFs on the likelihood of success for the task under consideration /Em84, vol.I,p.3/.

Assuming

$$x_i = 1 - y_i \quad (5)$$

this SLI model can be transformed into a model in which $\log p$ depends linearly on a failure likelihood index (FLI):

$$\log p = A + B \cdot FLI \quad (6)$$

where

$$FLI = \sum_{i=1}^n (w_i x_i) \quad (7)$$

$$0 \leq FLI \leq 1,$$

$$A = C + D,$$

$$B = -D.$$

It can be shown that eq. (1) (of the model proposed here) is equivalent to the FLI model in eq. (6) or to the SLI model in eq. (3):

$$\begin{aligned} \log p &= \log(p_o \prod_{i=1}^n K_i^{x_i}) \\ &= \log p_o + \sum_{i=1}^n (x_i \log K_i) \\ &= \log p_o + \sum_{i=1}^n (\log K_i) \sum_{i=1}^n \frac{x_i \log K_i}{\sum_{i=1}^n (\log K_i)} \\ &= A + B \sum_{i=1}^n (x_i w_i) \\ &= A + B \cdot FLI \end{aligned}$$

where

$$B = \sum_{i=1}^n (\log K_i) = \log \left(\prod_{i=1}^n K_i \right) \quad (8)$$

$$w_i = \frac{\log K_i}{\sum_{i=1}^n (\log K_i)} \quad (9)$$

$$A = \log p_o \quad (10)$$

The quantification procedure related to the FLI model in eq. (6) is outlined in Fig. 3.6; the single steps are described in the following sections. Fig. 3.6 shows two fundamental advantages of the equivalence between the equations (1), (3), and (6):

If the maximal factors ($K_i, i=1...n$) of the PSFs are known,

1. the PSF weights can be calculated via eq. (9), i.e., the laborious application of MAUT is not required, and
2. only one reference HEP is required to calculate the model parameters, A and B .

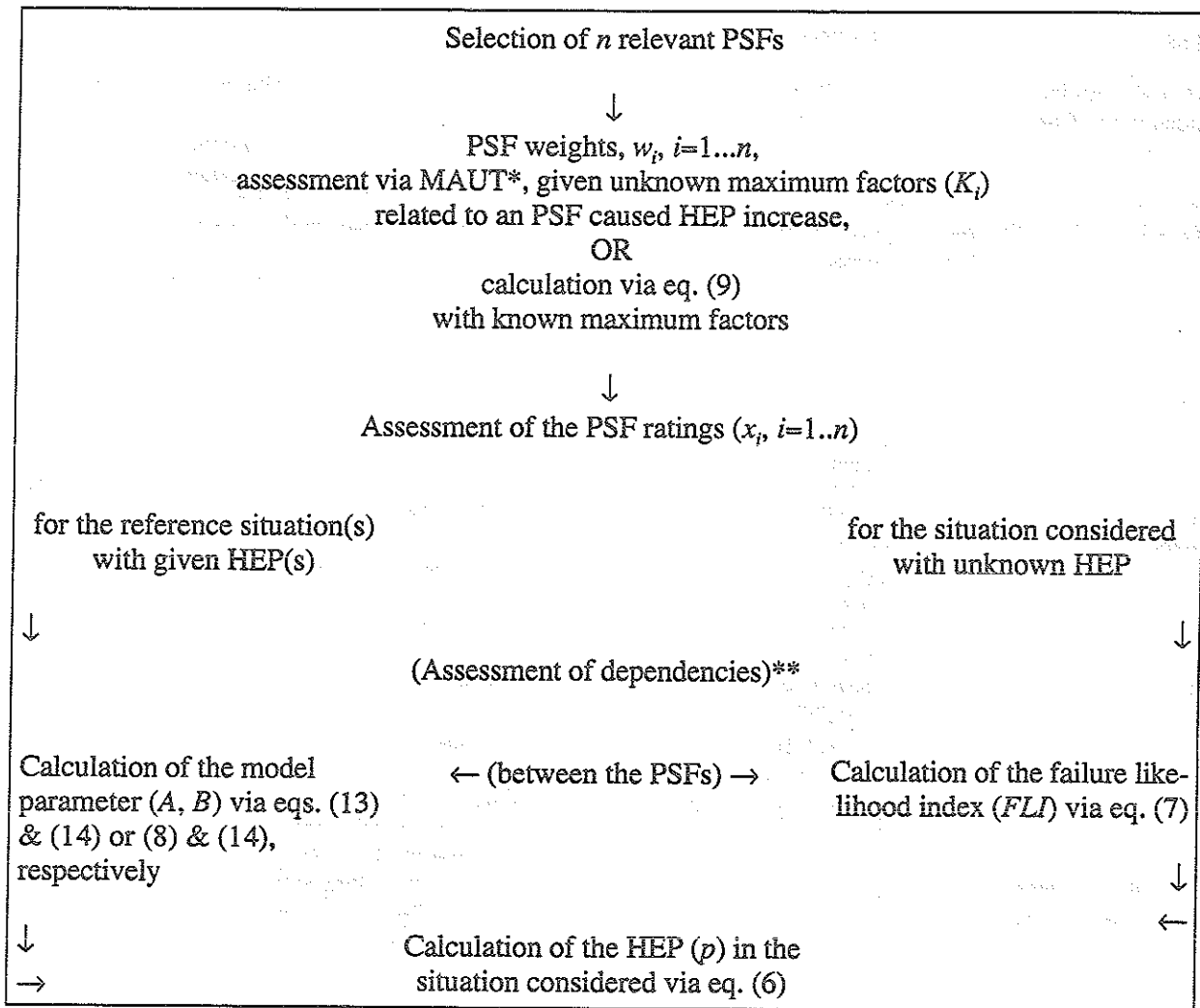


Fig.3.6: Method to assess the probability of decision-based errors in hazardous situations:

* Multi-Attribute Utility Theory;

** no component of the actual method, proposed field for further research.

3.4.3 Description of the PSFs selected to operationalize risk-taking behaviour

The selected PSFs related to risk-taking behaviour are listed in Tab. 3.4.

The selection of PSFs 1 to 5 was motivated by the classical approach of decision behaviour prediction (e.g., /Ed68/). These PSFs are related to the outcomes (costs, benefits, expected chances of occurrence) of safe and unsafe behaviour. The last two PSFs ($K_9=9$: unsafe behaviour too easy; $K_6=6$: safe behaviour unusual) are related to the real type of behaviour. These two PSFs can be summarized as "unsafe behaviour suggests itself".

The application of the PSFs listed above requires the estimation of x_i , the degree of presence. For PRA purposes, it is entirely adequate to represent a continuous variable by a finite number of points /Sw83, p.6-9/. Therefore, the method presented here proposes a five point discretization of x_i (see Tab. 3.5).

PSF	RATING PROPOSALS ***
($K_1=2$)*: An incentive to use other more dangerous procedures	($x_1=0$): Safe behaviour is only slightly inconvenient, e.g., using written procedure during work performance. ($x_1=1$): Safe behaviour is associated with a considerable disadvantage (e.g., plant shut down for days) which could be avoided by risk-taking behaviour.
($K_2=2.5$)*: A conflict between immediate and long-term objectives	($x_2=0$): No advantage for risk-taking behaviour exists; in this case the method proposed here is inappropriate. ($x_2=1$): There is some advantage that could be associated with risk-taking behaviour.
($K_3=12.5$)**: No personal safety affected	($x_3=0$): Due to risk-taking behaviour serious (safety-relevant) damage could occur which would be impossible, given safe behaviour. ($x_3=1$): Risk-taking behaviour could cause a problem of nonsignificant impact to safety.
($K_4=2$)*: Low failure expectation related to risk-taking behaviour	($x_4=0$): Relatively high chance that risk-taking behaviour will result in a negative outcome. ($x_4=1$): Relatively low chance that risk-taking behaviour will result in a negative outcome.
($K_5=4$)*: A mismatch between perceived and real risk	($x_5=0$): Based on interviews, the conclusion could be drawn that the operators have a realistic assessment concerning the frequency and consequences of possible failure events. ($x_5=1$): Based on interviews, the conclusion could be drawn that the operators have an over-optimistic realistic assessment concerning the frequency and consequences of possible failure events.
($K_6=9$)*: A means of suppressing or over-riding information or features which are too easily accessible	($x_6=0$): Risk-taking behaviour can be classified as a serious error of commission, e.g., an active intervention in the control of a safety system. ($x_6=1$): Risk-taking behaviour can be classified as an error of omission, and (in a later investigation of the related event) this behavioural error cause cannot be identified clearly, e.g., the operator could say that a misdiagnosis has caused this omission.
($K_7=6$)*: A need to unlearn a technique and apply one which requires the application of an opposing philosophy	($x_7=0$): Safe behaviour is not associated with a system state that contradicts the normal state. ($x_7=1$): Safe behaviour (e.g., initiation of feed & bleed cooling) is associated with a system state (e.g., primary circuit open) that essentially contradicts the normal state (e.g., primary circuit intact).
* /Wi88, EPCs 4,9,12,18,21/; ** /Sw80, pp.13-12 to 13-14/; *** rating effect: $PSF_i = (K_i)^{x_i}$	

Tab. 3.4: Selected PSFs for the probabilistic quantification of risk-taking behaviour

With given x_i estimations ($i=1...7$) to the PSFs presented in Tab. 3.4,

$$\underline{x} = (x_1, x_2, x_3, x_4, x_5, x_6, x_7) \quad (11)$$

the failure likelihood index (FLI) follows from eq. (7). The PSF weights follow from eq. (9) with input values from Tab. 3.4:

$$\begin{aligned} \underline{w} &= (w_1, w_2, w_3, w_4, w_5, w_6, w_7) \\ &= (0.068, 0.089, 0.248, 0.068, 0.14, 0.215, 0.176) \end{aligned} \quad (12)$$

For example, let

$$\underline{x} = (\text{high, total, no, unknown, unknown, high, total}) \quad (13)$$

be the ratings according to the non-initiation of feed and bleed cooling (as a type of risk-taking behaviour after total loss of feedwater), then the failure likelihood index is:

$$FLI = .068 \cdot .75 + .089 \cdot 1 + .248 \cdot 0 + .068 \cdot .5 + .14 \cdot .5 + .215 \cdot .75 + .176 \cdot 1 = .581 \quad (14)$$

NUMERICAL	LINGUISTIC
0	no presence
0.25	low likely presence
0.50	moderate likely presence OR unknown presence
0.75	high likely presence
1.0	total presence

Tab. 3.5: Discretization of PSF ratings

3.4.4 Using *one* reference HEP for calibration

Given an estimation for FLI , the calculation of the corresponding risk-taking behaviour probability p requires known values for the parameters A and B in eq. (6). In this sense, calibration means the method of deriving these parameters. The value of B follows from eq. (8); e.g., using the decadic logarithm:

$$B = \lg(2 \cdot 2.5 \cdot 12.5 \cdot 2 \cdot 4 \cdot 9 \cdot 6) = 4.43 \quad (15)$$

Firstly, the calculation of parameter A requires (at least) one reference probability (p_{ref}) relating to risk-taking behaviour. For this purpose, a median value of 0.01 is proposed. This value is taken from the most commonly accepted THERP data bank: a probability of 0.01 (based on industrial experience) that plant policies will be ignored is estimated /Sw83, p.6-18/.

Second, reference ratings (based on expert judgement) for each of the selected PSFs are required. These ratings are used to calculate the FLI value which corresponds to the reference HEP. Each reference PSF rating can be interpreted as the degree of presence which is already "included" in the HEP reference value of 0.01. An expert consensus (HRA team of the Research Centre Jülich) resulted in the following reference ratings for the PSFs in Tab. 3.4 corresponding to ignoring plant policies:

$$\begin{aligned} \underline{x}_{ref} &= (x_{1ref}, x_{2ref}, x_{3ref}, x_{4ref}, x_{5ref}, x_{6ref}, x_{7ref}) \\ &= (\text{no}, \text{total}, \text{high}, \text{unknown}, \text{low}, \text{high}, \text{low}) \end{aligned} \quad (16)$$

Using the weights in eq. (12), the ratings above result in a reference FLI value as follows:

$$FLI_{ref} = .068 \cdot 0 + .089 \cdot 1 + .248 \cdot .75 + .068 \cdot .5 + .14 \cdot .25 + .215 \cdot .75 + .176 \cdot .25 = .548 \quad (17)$$

Finally, parameter A follows from the eqs. (6), (15), and (17):

$$A = \lg p_{ref} - B \cdot FLI_{ref} - \lg 0.01 - 4.43 \cdot 0.548 = -4.43 \quad (18)$$

The value of parameter A depends strongly on two components of expert judgement: (1) the selection of a suitable reference HEP and (2) the corresponding PSF ratings. Therefore, it is important that these judgements result from a consensus of experts who are familiar with HRA methods concerning cognitive aspects of human behaviour.

The PSF ratings corresponding to the unknown HEP (eq. (11)) and the PSF ratings corresponding to the known reference HEP (eq. (16)) should be estimated by the same HRA

team. In doing so, the HEP calculated by eq. (6) could be interpreted as a value which lies - from the subjective view of the judging HRA team - in a fair relation to the known reference HEP. Eq. (19) demonstrates this coherence quality:

$$p = p_{ref} \prod_{i=1}^n K_i^{x_i - x_{ref(i)}} \quad (19)$$

Note: eq. (19) follows from eq. (1) by substituting p_0 by p_{ref} using eqs. (18), and (7) to (10).

3.4.5 Uncertainties

The model includes the uncertainty quantification of the basic HEP (p_0) and of the adjusting factors (PSF_i , $i=1\dots n$). Point estimates of probabilities or factors are interpreted as median values. The uncertainty of a factor is reflected by the uncertainty of the related rating (x_i). Referring to the distance (0.25) between two discrete points of x_i , the 90% uncertainty interval is assumed to be expressed by $x_i \pm 0.25$. Given an estimate of UNKNOWN, an interval of 0.5 ± 0.5 is assumed. Moreover, it is assumed that the uncertainty of a rating is reflected by a normal distribution. Given this assumption, the uncertainty of a factor (PSF_i) is reflected by a lognormal distribution with an error factor of

$$EF_i = K_i^{0.50}, \text{ given } x_i = \text{unknown},$$

$$EF_i = K_i^{0.25}, \text{ otherwise,}$$

where K_i is the maximum value of PSF_i .

For example, given the estimates in eq. (16) for the ratings of the reference HEP ($MDN_{ref}=0.01$, $UF_{ref}=5$), the error factor of the basic HEP is:

$$UF_o = \exp \left\{ \left[(\ln UF_{ref})^2 + (\ln UF_F)^2 \right]^{0.5} \right\} \cong 7$$

where

$$UF_F = \sum_{i=1}^7 (\ln UF_i)^2.$$

Consequently, the constant of the FLI model,

$$A = \lg p_o = -4.43,$$

is associated with an uncertainty of

$$\Delta A = \pm \lg 7 \cong \pm 0.85.$$

3.4.6 Related software

The quantification of decision-based errors is supported by a small computer code. It employs functions of the "Statistic Analysis System" software /SAS89/. However, the calculations are quite simple. Hence, computer support is not implicitly required.

3.4.7 Future research fields

The model presented here is an initial approach. A calculated HEP reflects the opinion of the model user about the ratings of seven PSFs (in the situation to be analysed) compared to the ratings in a reference situation (with known HEP).

One future research field could concern the coherence of such opinions, i.e. the reduction of the user-caused result variability; see Tab. 3.6. This could be realized by benchmark studies /Po88/. For example, a high user caused variability shown by such studies could be reduced by an improved description of the PSFs listed in Tab. 3.4.

Another research field could concern the predictive accuracy: does the model give an HEP prediction which is - with respect to the usual PRA uncertainties (i.e. $EF=3...10$) - quite accurate? However, without data, the accuracy improvement still depends on expert judgement. Tab. 3.6 outlines the use of expert judgement for model validation. For example, the model could be extended by a submodel with respect to a plausible interaction between two PSFs, or, as a result of a pairwise comparison, the PSF weights could be adjusted.

Nevertheless, the data problem could produce a positive accompanying phenomenon: PRA practitioners need support from experts in psychology and ergonomics. Cooperation between different disciplines could be based on this resulting in positive outcomes for all concerned.

OBJECTIVE	REALIZATION	POSSIBLE OUTCOME
Reduction of result variability • given several users at the same time • given one user at several times	Benchmark studies	Rating support • improving descriptions of PSFs and reference points • adding reference points
Improving predictive accuracy	Validation studies • using empirical HEPs • using expert judgement, e.g. background knowledge about dependences, pairwise comparison related to the relative importance of PSFs, etc.	Model adjustments • changing PSF weights • adding a PSF • taking a PSF away • including PSF interactions

Tab. 3.6: Future research fields concerning a model for the probabilistic quantification of risk-taking behaviour

3.5 Software supporting human reliability analysis and evaluation

3.5.1 Concept of software system

As was described, human reliability analyses are complex tasks which might be performed more effectively using special software to support the HEP evaluation and documenting the analyses. Such software can be designed, in its initial version, based upon conventional programming techniques (including access to data bases) and then evolved into more advanced software systems using the expert system technology and AI techniques. The documenting of the analyses and evaluations is important for future scrutinizing of the HRA/PSA results. Thus, three stages in developing such software can be proposed:

- (1) Conventional software for the human reliability evaluation and documenting;
- (2) Applying the expert system technology for human reliability evaluation and documenting based upon several HRA methods;
- (3) Software for advanced modelling of process systems and cognitive operator behaviour in dynamic event-driven situations for supporting HRA.

The main design features of these three respective types of software are as follows:

- (1) HRA is menu-driven to permit the expert to select a desired method from the techniques available and feed relevant attributes. Then HEP is calculated and stored in a data base. The content of the data base can be modified. When required the graphical option of HRAET or OAT is activated to construct an HRA logical model with support from the probabilistic calculator. For multi-expert methods a special panel is activated for acquiring expert opinions and a procedure for aggregation of estimates. The results of the human error probability evaluation are documented in the plant-specific human reliability data base (HRDB) in which human failure events are depicted using special coding symbols. The software as described above can support the conventional PSA. An example of similar software is OCRA described in Item 2.3.6.
- (2) The software based on expert system technology contains expertise which usually evolves in time according to the state of art of HRA and new techniques developed. The system should be designed to help the user to select the appropriate HRA method or methods for the case analysed (selection of techniques can be based on classification trees presented in Item 3.2). Searching in data base (e.g. of the THERP technique) and evaluation of HEPs are performed in a semi-automatic way according to heuristics and a model with input parameters associated with the given technique. Results of evaluation are documented in a plant-(project)-specific HRDB. The software described can be useful for performing either the conventional PSA or PSA based on the expert system technology (Section 4.2). An important part of such a software system could be an expert system for the application of THERP described in the next item.
- (3) Such software might be designed for the qualitative evaluation of operator behaviour in dynamic event-driven accident conditions with regard to the window time for operator response, important PSFs (e.g. stress), the training scope of operators, the quality of operating procedures and the design of man-machine interface, etc. The purpose is to investigate possible situations that can aggravate the course of accident and human error intentions due to a coincidence of various factors that cannot be discovered in the conventional HRA/PSA study. Results of such analyses would be useful not only for HRA/PSA experts in verifying the event trees (new significant accident sequences), but also of paramount importance for: training purposes, designing of the man-machine interface, verifying of procedures, classifying the human behaviour types (rule or knowledge), assessing the quality of the plant interface, generalized assessing of the plant states and the

management of the plant. The software system possessing mentioned features might be designed in the future for more abstractive modelling of the plant responses and human behaviour using AI methods with regard to human behavioural sciences and the psychology of errors.

In most current HRA studies the results from indirect deterministic simulations are taken for evaluation of the allowable response time required as an input to the e.g. HRC or THERP techniques. The direct simulation of the plant would be preferred, e.g. for evaluation of similarity measures of the plant responses for different initiating events to support quantitative probabilistic evaluation of the diagnosis error, as was described for the CM-MS method /Ko92/.

An interesting approach is that applying the DYLAM methodology described in Item 2.5. A simplified deterministic simulation tool would be useful for developing object-oriented models of the plant, which are, in their simplest form, relations on the plant variables /Wo88,Ca89/. Such representation of the plant dynamic properties as an abstraction requires further research to be applied within the mentioned AI modelling system.

However, as described in Item 2.5 these new approaches support only the analysis of human behaviour at higher cognitive levels (e.g. identifying of intention errors) and are at present not directly useful for quantifying of the human reliability. For that purpose the multi-expert techniques can be applied, e.g. the SLIM method /Wo88/. The new human behaviour analysis techniques should be verified and validated before using them in engineering practice, e.g. with regard to results obtained from experiments on the dynamic simulator.

The consideration of the practical needs for quantifying of HEPs (see Item 3.2 and 3.3) led us to the conclusion that at present the most urgent need is the development of a computerized version of the THERP technique. It was decided to develop the relevant software based upon the expert system technology. This system might in the future be a part of the more complex HRA expert system which would consist of several human reliability modelling techniques.

3.5.2 An expert system for the application of THERP

The analysis and evaluation of safety-relevant operator errors is an essential component of probabilistic risk assessments. In practical application, the high scattering of evaluation results obtained by different analysers appears problematic. Improvements can be achieved by avoiding errors in the application of methods, ensuring completeness of method application, at least to a large extent, and disclosing subjective influences. In all three respects the ESAP expert system /Re93/ will enhance the application of the THERP methodology.

THERP was selected because it is a well-documented method suitable for evaluating human errors prior to and during an accident, and thus comprehensive. THERP is commonly accepted among PSA users; the required data are contained in the associated handbook.

The setup and functioning of ESAP are described in Appendix 4.

3.6 Main steps for further system development

The next stage of the system development should be the design of the HRA expert system based upon several techniques to cover the important cases of human interactions/errors discussed in Section 3.1 and 3.2. Taking into account candidate modelling techniques for the three levels of the HRA/PSA effort (Fig. 3.4 and 3.5) it is proposed to develop an HRA expert system supporting:

- (1) Selection of the modelling technique for different kind of interactions and for mode and effect of actions/errors (Fig. 3.4 and 3.5);
- (2) Evaluation of HEPs for relevant attributes/parameters for the model selected;
- (3) Documenting HRA and results of the HEP evaluation for single- and multi-expert techniques.

After implementing the THERP technique it is proposed to develop a computerized version of the following techniques: ASEP-HRAP, INTENT, SLIM and APJ as well as an HRA co-ordinating module using the expert system technology. The software system should be developed in such a way as to enable operation within the data base environment. This HRA expert system would then be part of a PSA expert system described in Item 4.2.

4 Integration of computer-aided HRA with PSA

4.1 Incorporation of human-induced events into PSA logic

The conventional PSA studies are performed by experts who use supporting computer programmes. The logic of the fault/event trees is developed by the expert(s). The computer programmes are used mainly as tools for their graphical representation and probabilistic evaluation, sometimes also with access to data bases or supporting calculations, e.g. for the human error probability evaluation. Thus, the effort of the logic development using the fault/event trees and probabilistic modelling is exerted by experts. The conventional PSA can be supported by a special software for the human reliability evaluation based upon the conventional computer programme or the expert system technology.

In the first stage of modelling the fault/event trees are constructed by the system analysts and are in principle system-oriented. In the next phase they are verified and/or modified by human reliability analysts to include failure events induced by human factor. A widely used framework to perform HRA and incorporate the human-system interactions into PSA is Systematic Human Action Reliability Procedure (SHARP) /Ha84/. A similar framework is proposed by the IAEA (see Section 2.4). Adapted frameworks should include the human failure events classification proposed in Section 3.2 with regard to techniques of the HEP evaluation corresponding to HRA/PSA effort (low, medium or high).

Since the effects of slips during plan-directed activities do not usually propagate unassisted by other failures beyond the component(s) affected, they can be incorporated into fault tree models. Mistakes, which are most significant in event-driven activities, are best incorporated at the top of or above the system fault tree level because of their propagative potential. Recovery events are tacked onto the end of the dominant sequences to which they apply and may not be formally included in the event or fault trees (they can be left as adjoints to cut sets). In this way, the classification effort supports the integration effort. Levels of incorporation of human failure events in PRA models are illustrated in Fig. 4.1. In the HRA/PSA screening process only more probable human failure events are taken into account for further consideration /Ha84/.

The following guidelines can be formulated for incorporating human-induced failures into the hardware-oriented logic /Do88/:

- (1) The systems analysts should only describe human failure events functionally (not too specifically).
- (2) Only three types of human failure events should be identified - failure before an initiator, called latent; human-induced initiators; and post-initiator human failures.
- (3) All latent human failure events should be incorporated into the appropriate fault tree at the highest level that accounts for all of the equipment manipulated during the test or maintenance task. The highest level is typically the train or part-train level. A task is the sum of all the actions that lead to a test or maintenance objective, including system isolation, the intended activity, and system restoration.
- (4) Human failures that lead to an initiator should be assumed to be included in the data for the initiator class that includes the human failure and not modelled explicitly. In the case that a special hardware initiator is modelled, the fault tree can include latent error events identified according to guideline (2).
- (5) All post-initiator human failure events that relate to the manual actuation of a system should be modelled as a single event for that system in an "or" gate at the top of the system

model. The project may alternatively model such an event in the top logic of the appropriate event tree. A type of post-initiator human failure event is sometimes referred to as a commission error (e.g. operators operate system Y rather than X). Such events should not be routinely located in the logic structure, unless a special situation is noted that could give credence to such an event.

- (6) Where a human action is required for the system actuation, for example, manual initiation of scram, a corresponding human failure event is logically associated in an "and" gate with the failure event of the system's actuation. These events are often termed recovery events and amount to failing to restore the function of faulted equipment or failure to find alternatives to its function. Such recoveries may be identified and described early in the systems analysis but should be not modelled until the total sequence cut set has been identified. At this point a recovery analysis will identify the appropriate recovery actions, given the sequence context.
- (7) The HRA needs supplementary information from the system analysts to quantify any human failure events identified by the systems analysts.

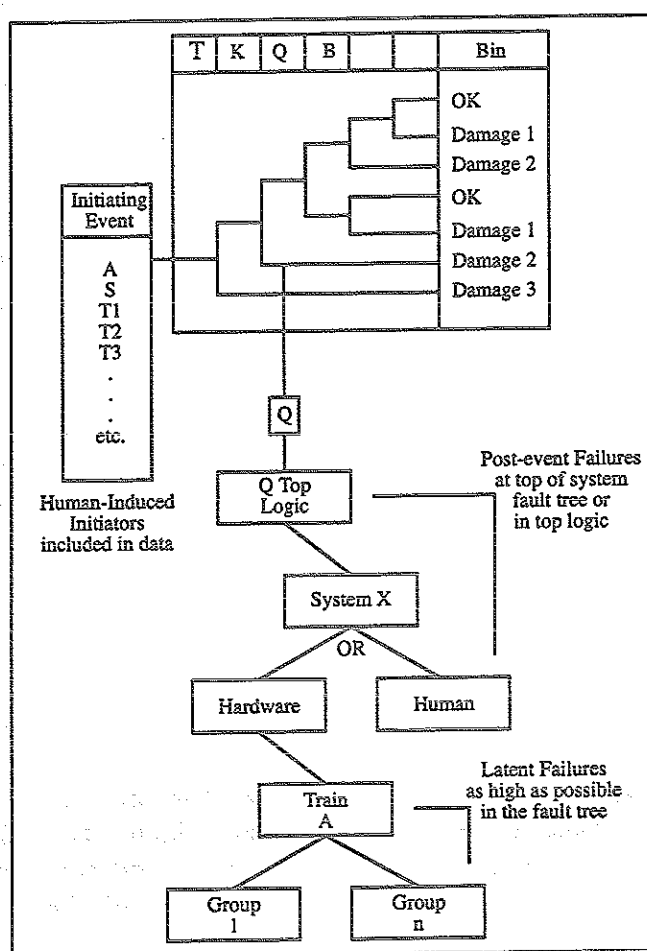


Fig 4.1: Levels of incorporation of human failure events in PRA models /Do88/

In this approach the experts are responsible for selecting appropriate techniques for human failures evaluation and other tasks described at the beginning of this section. Most of these tasks can be performed by an expert system having built-in expertise acquired from HRA experts.

Thus, the reliability models of human factor and probabilistic models of the plant's equipment are to be integrated at different levels of hierarchy: components, subsystems and systems. An example of including recovery events in the fault tree is illustrated in Fig. 4.2. Recovery events concerning the higher cognitive level of the operator behaviour can be included in events trees /IA91c, IA92b/. Including recovery events in the logic models requires careful consideration to limit the calculation cost of quantitative evaluation of the entire system /Do88/.

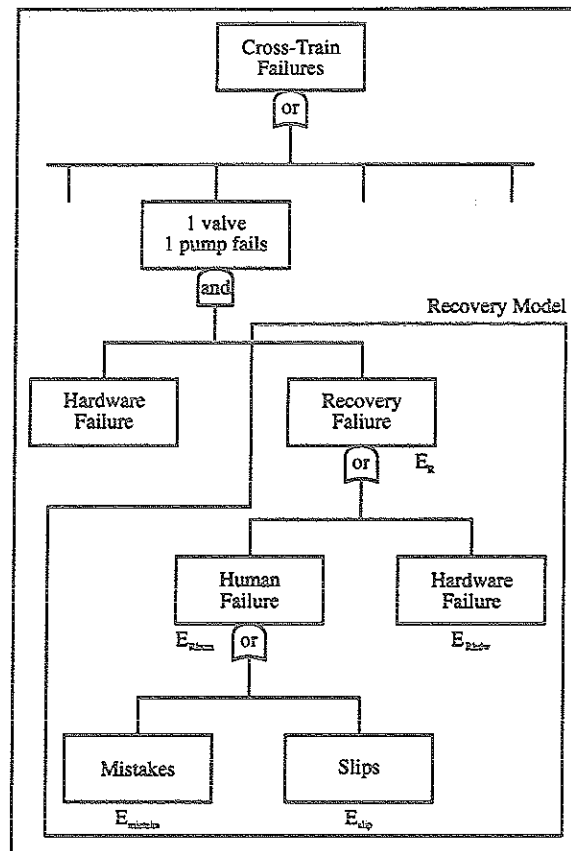


Fig. 4.2: Recovery fault tree model for recirculation failure event E /Do88/

4.2 Application of the ES technology for HRA/PSA

Several prototype expert systems supporting PSA are described in Appendix 2. These expert systems are in principle system (hardware) modelling oriented. Functions and processing phases of the PSA/HRA expert system proposed to be developed are presented in Fig. 4.3. The system should be designed for PSA/HRA experts to facilitate Probabilistic Safety Analyses of complex technological systems, nuclear power plants in particular. This system will be also very useful for training purposes thanks to explanation components and menu-driven procedures. The objective is to perform Probabilistic Safety Analyses effectively in a semi-automatic way.

The system should be designed to support "living" PSA with a contribution by PSA experts. Four general tasks/functions can be distinguished:

- A. General data/knowledge acquisition.
- B. Plant-specific data/knowledge acquisition.
- C. Inferring/evaluation process supervised by the expert/user.
- D. Presenting and/or documenting of data/knowledge and results of the analyses/evaluations.

The system will support the performance of probabilistic safety analyses supervised by the expert/user in the following phases:

- (1) Input of topological data concerning the safety front line and supporting systems with technical parameters and other attributes;
- (2) Input of functional/logical data and knowledge associated with these systems with the attributes of components;
- (3) Semi-automatic event tree construction;
- (4) Semi-automatic fault tree construction;
- (5) Modification of fault/event trees with regard to human actions and errors;
- (6) Modification of fault trees regarding dependent failures of systems;
- (7) Modifying or creating some event/fault trees by the expert or user;
- (8) Determination of minimal cut sets (MCSs) for top events of fault trees and for accident sequences with relevant Boolean reduction of events;
- (9) Quantifying of MCSs, top events and accident sequences;
- (10) Sensitivity and reliability importance analysis;
- (11) Uncertainty assessment;
- (12) Documenting the analyses and evaluations.

To accomplish these tasks the system will consist of several data/knowledge bases. Their development requires a considerable initial effort. The logical and quantification process leading to final results will be supervised by the expert/user. An important part of the system is the explanation and decision support module (Fig.4.3). The expert/user will be able to modify the topological/functional data as well as the hardware-oriented and human reliability models with relevant input data to investigate their influences on final probabilistic results. The features and the design phases of a software system using the expert system technology are described in Appendix 1.

To accomplish these functions three scopes of the software system development can be considered which are characterized in Table 4.1, related to the PSA/HRA effort, respectively 'small', 'medium' and 'high'.

The small effort is an initial HRA/PRA in which consecutive tasks are performed usually by a single user/expert. The analyses are rather equipment-oriented with moderate attention paid to

common cause failures, human failure events and the recovery actions. The probabilistic assessments are based on the point estimators with less attention paid to uncertainties. If required, the reliability importance analyses are performed and finding the candidates for most important accident sequences for further more extensive evaluation. Construction of the event trees is not supported directly by deterministic modelling and simulation. The experts rather possess only qualitative knowledge concerning the plant response for various initiating events and for assumed scope of functional failures within safety systems.

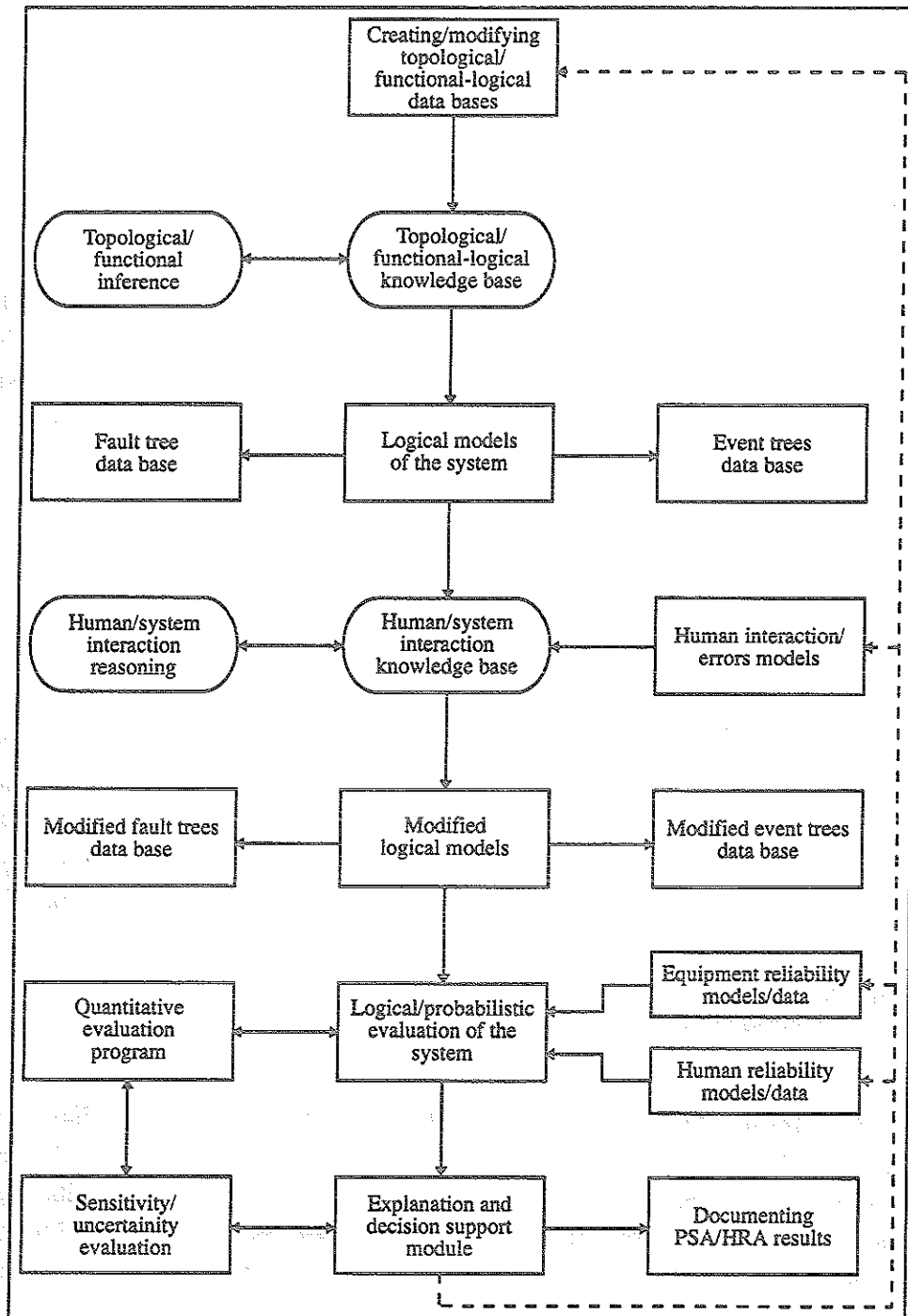


Fig. 4.3: Functions and processing phases of a PSA/HRA expert system

Description of features of PSA1/HRA and supporting software	Scope I *	Scope II *	Scope III *
Exhaustiveness of analyses	Low	Moderate	High
Cost of assessments	Low	Moderate	High
Min. users (Us)/ experts (Es) involved in analyses	Single U	Single E	Several Es
Menu-driven support and explanation facility	Extensive	Moderate	Moderate
Heuristics	Simple heuristics	Detailed heuristics	Consensus heuristics
Access to results of deterministic simulation codes	From reports	Indirect	Direct
Object-oriented modelling	Only a few logical models	Some	Yes
Human reliability models	Fig. 3.4/I Fig. 3.5/I	Fig. 3.4/II Fig. 3.5/II	Fig. 3.4/III Fig. 3.5/III
Dependent failure analyses	Simplified	Exhaustive	Exhaustive
Uncertainty treating framework	Bayesian Probability or Fuzzy Set Theory	Bayesian Probability or Fuzzy Set Theory	Bayesian Probability or Fuzzy Set Theory
Details in Topological Data Base	Moderate	Moderate	Yes
Semi-automatic fault/event trees construction	Some	Most	Yes
CAD graphical support	Yes	Yes	Yes
Interface to external files and data bases	Yes	Yes	Yes
Hardware	PC-386	PC-486	Work station
Use for training	Initial HRA/PSA training	HRA/PSA training	Advanced HRA/PSA training/research
Use for reliability and safety-related decision making	Initial decisions	Preliminary decisions	Detailed analyses and final decisions

Tab. 4.1: Assumed features of software supporting PSA-level-1 and HRA based on the expert system technology

* Related to the effort defined by Figures 3.4 and 3.5

The medium effort HRA/PSA is extended to include more extensive analysis of human failures and common cause failures. Possible recovery actions are investigated in detail. The available reliability data bases and other various relevant information sources are taken into account. The uncertainty bounds of the probabilistic results can be assessed as fuzzy (see Section 4.4 - Bayesian framework is more restrictive). Constructing event trees can be supported by simplified deterministic modelling and simulation (indirect access) with classifying of the plant states in emergency situations. The plant states and accident sequences are pointed out which require further multiple expert analyses and evaluation.

The high effort HRA/PSA is performed with contributions by several experts to decide on the most important and controversial modelling cases. Probabilities of the human errors are assessed with more detailed analyses of PSFs. Frequencies of rare hardware failure events and dependent failures are assessed using multi-expert evaluation techniques. Two approaches to represent uncertainties are proposed: based on fuzzy measures and Bayesian measures. The second approach is justified only in cases of hard input probabilistic data (Section 4.4). Because the expert system with complex heuristics represented by rules is difficult to maintain it is proposed for the scope III of the system development (Table 4.1) to apply object-oriented modelling more extensively with relevant relations concerning logical modelling of the safety systems and supporting deterministic modelling of the plant. As reference codes the conventional simulation codes might be employed, for this PSA effort /Me89/, run e.g. on a PC parallel processing simulation system /Bi92/.

4.3 Hardware and software environment for advanced HRA/PSA

Requirements for the hardware and software environment for the development of the HRA/PSA expert system depend on the scope of the PSA/HRA studies (Table 4.1). The software environment for building complex expert systems can consist of the modules shown in Appendix 1 (Fig. A1.1, A1.4 and A1.5). In general such software is designed using the expert systems shells, general-purpose programming languages, higher-level programming languages, software development kits and other supporting software (Fig. A1.4). Expert systems and other AI technologies can be added to the DSS (Decision Support System) consisting of the data and model bases with relevant management software facilities (Fig. A1.5). There is extensive literature available concerning expert systems and other applications of AI techniques, e.g. /Da79, Da80, Ha83, Ha85a, Ni86, Le86, Ro88/.

The software system for advanced PSA(level 1)/HRA studies should consist of the following components:

- (1) Hybrid expert system shell (and an AI supporting software)
- (2) CAD (Computer Aided Design) support for graphical representing of topological information
- (3) Data base environment and its management compatible with other software applied
- (4) Model base environment and its management supporting modules
- (5) Interface to the user.

The system should be designed in such a way as to enable rapid prototyping of new expert system functions. A good starting point is to decide about a shell for building the expert system. Possible candidates are KEE (Appendix 2), EGERIA, PROKAPPA and in the future KAPPA 3.0 (Appendix 1). The system selected should enable expert systems to be designed as the blackboard architecture and possess possibilities to reason under uncertainties (see next Section).

The most difficult and costly aspect will be the development of the software supporting PSA/HRA of scope III (Table 4.1), especially as regards an interface to deterministic models. This problem should be considered carefully due to the great research and designing resources required. As was discussed in Item 3.5 two cases can be considered:

- (1) Human reliability modelling using conventional HRA methods.
- (2) Human behaviour and error intention analyses in complex dynamic man-machine system using an approach based on AI methods.

For the first case the conventional modelling and simulation codes can be applied, as e.g. MELCOR /Bi92/ or simplified deterministic simulation programmes (as e.g. DSNP).

It seems useful to support the second case by different computer tools. Modelling of the process system is performed at a certain level of abstraction. The relevant research should be aimed at object-oriented modelling and applying AI methods. Good starting points for further research concerning modelling of man-machine interface are CES and COSIMO methods described in Section 2.5. Thus, the methodology is needed to develop an hierarchical object-(event-) oriented model of the plant based upon expert knowledge supported by a deterministic simulation code (applying learning capabilities of AI systems can be considered) which will be compatible with the human behaviour model. Because such models do not produce the HEPs, the development of a method will be required to transfer findings of HRA to the probabilities using expert opinions.

4.4 Methodological issues for treating uncertainty in HRA/PSA

4.4.1 Bayesian and non-Bayesian theories

In the quantification process of the fault and event trees the relevant probabilistic models of systems including equipment and human failures are combined to obtain the final results (probabilities of the fault trees top events and accident sequences) with associated uncertainty measures. Thus, more or less "hard" and "soft" probabilistic data and intermediate results are combined at different levels of the model hierarchy to obtain the overall final probabilistic results. In this process the recovery events should be included in such a way as to limit the number of minimal cut sets (MCSs) generated and to reduce the calculation cost/Do88/. Several frameworks for uncertainty representing and treating can be considered /Ka86a, Du88a, Wu90/:

Bayesian probability approach

Current PSA methodology extensively adopts the subjective probabilists' point of view, in which probability is the measure of degree of belief in the truth of an event, based on one's current state of knowledge /Wu90/. The standard mathematical theory of probability is consistent with such an interpretation. The foundations of the mathematical theory of probability are the following axioms:

Let 2^S denote the set of all subsets of a sample space S related to define events. A probability function, $p(A)$, associated with each event in 2^S is a real number satisfying the following principles:

- (1) $p(A) \geq 0$ for each event A
- (2) $p(S) = 1$
- (3) If there exist a set of mutually exclusive events $\{A_1, \dots, A_n\}$, then

$$p(A_1 \cup A_2 \cup \dots \cup A_n) = p(A_1) + p(A_2) + \dots + p(A_n)$$

Thus, an essential property of a probability function is the additivity of the probabilities of mutually exclusive events. The assessment of probabilities that complies with this axioms is said to be coherent. This requires, for instance, that if the probability p is assigned to an event then probability $(1-p)$ ought to be assigned to the complement of this event.

Since a large portion of PSA deals with state-of-knowledge uncertainty, the modelling of the degree of belief has a direct impact on PSA methodology. It has been pointed out that certain uncertainties associated with the event trees are rather of the state-of-knowledge type than of a stochastic nature /Wu90/: only one of several hypotheses can be true, and one must express one's degree of belief in the truth of these hypotheses.

In the Bayesian framework only one piece of information, regardless of the difficulty of obtaining the numerical values, the subjective probability, p , is needed for an event A (one of the hypothesis). The degree of belief that an event A will occur is p , and the degree of belief that the event A will not occur is, by the third axiom, $(1-p)$.

Another model can be applied allowing relaxation from this axiom which would require, e.g., two separate pieces of information: the degree of belief that an event A will occur, and the

degree of belief that event A will not occur. Two other models will be considered, based on the theory of evidence and the theory of possibility.

Belief functions

Shafer introduces, instead of a probability function, the concept of belief function in his theory of evidence /Sh76/ as the model for representing knowledge under uncertainty. The belief function is defined as follows:

Suppose that S is a finite set, and let 2^S denote the set of all subsets of S. A belief function, associated with each event A in 2^S is a real number between 0 and 1 such that the following conditions are true:

$$(1) \text{Bel}(\emptyset) = 0$$

$$(2) \text{Bel}(S) = 1$$

(3) For every positive integer n and every collection $A_1, \dots, A_n$ of subsets of S,

$$\text{Bel}(A_1 \cup \dots \cup A_n) \geq \sum_i \text{Bel}(A_i) - \sum_{i < j} \text{Bel}(A_i \cup A_j) + (-1)^{n+1} \text{Bel}(A_1 \cap \dots \cap A_n)$$

The rule 3 for a degree of belief for the disjunction of events is a relaxation of the principles of the Bayesian approach probability. A direct consequence of this relaxation is that the sum of the degrees of belief in an event and its complement does not need to be equal to certainty. Mathematically it follows that:

$$\text{Bel}(A \cup \bar{A}) = 1 \geq \text{Bel}(A) + \text{Bel}(\bar{A})$$

or

$$\text{Bel}(A) \leq 1 - \text{Bel}(\bar{A})$$

The belief function of an event A reflects all the knowledge gained from the evidence that supports that event. A different, but related measure is the plausibility function, $Pl(A)$, defined to reflect the knowledge gained from the evidence that does not support $\bar{A}$, i.e.

$$Pl(A) = 1 - \text{Bel}(\bar{A})$$

The belief function always has a numerical value lower than the plausibility function. In fact, they are originally derived and referred to as the upper and lower probabilities of a given event in Dempster's paper /De67, De68/.

In order to combine the independent pieces of evidence or sources of information this theory adopts Dempster's rule of combination /Sh76/. However, when pieces of information are in conflict this rule requires that the evidence be removed or ignored, and therefore cannot be applied in its basic form /Za86/.

Possibility

The theory of possibility is developed based on the concept of fuzzy logic /Za78, Za83, Za86, Ga76/. Its treatment of state-of-knowledge uncertainty can be viewed as equivalent to a relaxation of axiom 3 of the probability theory. Instead of the property of additivity, the theory of possibility applies a weaker operation to the disjunction of multiple events. The possibility function can be defined as follows:

Suppose that S is a finite set and let 2^S denote the set of all subsets of S . If function Π maps the set 2^S on a real number interval $(0,1)$ and satisfies the following conditions:

$$(1) \Pi(\emptyset) = 0$$

$$(2) \Pi(S) = 1$$

(3) For every positive integer n and every collection $A_1, \dots, A_n$ of subsets of S ,

$$\Pi(A_1 \cup \dots \cup A_n) = \max \{ \Pi(A_1), \dots, \Pi(A_n) \}$$

then Π is called a possibility function over S .

When the maximum is chosen as the operator for the disjunction of events, the events being mutually exclusive is no longer needed as a sufficient condition for the statement. Conditions 2 and 3 lead to the following:

$$\Pi(A \cup \bar{A}) = 1 = \max \{ \Pi(A), \Pi(\bar{A}) \}$$

In other words, knowing the possibility of an event A does not inform us fully about the impossibility of its complement. The necessity function, $N(A)$, is therefore defined to reflect the belief in the impossibility of $\bar{A}$:

$$N(A) = 1 - \Pi(\bar{A})$$

It follows from above that

$$N(A) \leq \Pi(A)$$

and

$$\max \{ \Pi(A), 1 - N(A) \} = 1$$

It is pointed out that possibility and necessity can be viewed as another form of upper and lower bounds on an unknown probability value. Any possibility distribution is a fuzzy set membership function /Du88a, Du88b/. It can be shown that formulas as above lead to wide upper and lower probability bounds which provide little information /Wu90/. However, the literature on the application of fuzzy mathematics to various technical fields is vast and growing /Ka86a, Ka86b, Ka86c, Be87, Du88a, Du88b/.

The foundations of the theories described above and the consequences of their applications in assessment practice and decision making have lately been issues in many scientific disputes. Each of these theories has its advocates and opponents /Ka86a, Le88/. There are many advocates of the traditional theory of Bayesian probability. There is, however, growing interest in the application of new theories (the evidence theory and the possibility theory) in practice, e.g. for representing uncertainty and the evidence combining in a new computer application based on AI methods and the expert systems technology.

4.4.2 Combining knowledge

The PSA results depend heavily on expert judgement /Ap88b, Hu88, Mo88/. An example of the PSA study is NUREG-1150 /Nu87/, in which expert opinions have been extensively applied. There is, therefore, a need to devote more research to this issue and to develop a relevant framework to combine expert opinions within HRA/PSA. Basic theories which can be considered are described below.

Combining knowledge in the Bayesian framework

The fundamental tool for combining information from several sources (or the incorporation of new evidence into an existing body of knowledge) in the Bayesian probabilistic framework is Bayes' theorem. This framework has won considerable interest within the PSA research community /Ap89, Wu90/. Its use for combining information from different information sources, also dependent sources, including experts has been the subject of research in management science /Mo77, Wi81/. The Bayesian framework is relatively well known and it will be not presented here further in detail, but it will be discussed latter in the context of other frameworks for combining information from different sources including experts.

Combining knowledge in the theory of evidence

The rejection of the additivity of the probabilities of mutually exclusive events invalidates Bayes' theorem /Wu90/. The theory of evidence can be seen as a generalization of Bayesian probability concept /Sh76/. Within the theory of evidence for combining pieces of evidence Dempster's rule of combination is used. Since expert opinions are often treated in PSA as pieces of evidence, it has been suggested that Dempster's rule of combination can also be used for combining expert opinions /Ka86a, Ga87b, Le88/. The application of this framework has also been proposed for risk assessment /Ho89/.

Shafer's idea is to partition one's belief among the different subsets of 2^S assigning to each subset B the portion of total belief that is committed strictly to B and cannot be further partitioned among its smaller subsets. The result is the set of so-called focal elements $\{B/m(B)>0\}$, m is called the basic probability assignment (or basic probability mass). A mapping m allocates to each event a portion of the total knowledge, i.e. the sum of all basic probability masses in 2^S must be equal to 1 and for an empty set $m(\emptyset) = 0$. The degree of belief in some event A is then calculated on the basis of events whose occurrence imply that of A , i.e.

$$Bel(A) = \sum_{B \subseteq A} m(B)$$

The plausibility of A is to be evaluated on the basis of sets B for which $B \cap A \neq \emptyset$

$$Pl(A) = \sum_{B \cap A \neq \emptyset} m(B)$$

As was mentioned the probability of an event A , $P(A)$, is related to the belief and plausibility by:

$$Bel(A) \leq P(A) \leq Pl(A)$$

For two independent but interacting pieces of evidence with basic probability assignments m_1 and m_2 , consisting of focal elements $A_1, \dots, A_k$ and $B_1, \dots, B_n$ respectively, a combined probability assignment is given by the orthogonal sum:

$$m_{12}(A) = (1 - k)^{-1} \sum_{A_i \cap B_j = A} m_1(A_i) m_2(B_j)$$

where

$$k = \sum_{A_i \cap B_j = \emptyset} m_1(A_i) m_2(B_j)$$

is related to the degree of conflict between these two structures; $k=0$ for the consensory sources of information. For conflicting sources of information ($k=1$) this formula cannot be applied. In Dempster's rule of combination it is not possible to include the credibility of various sources of information, however, some modifications of this rule are possible based on some ideas supported by the fuzzy set theory /Du88a, Du88b/.

Combining knowledge in the theory of possibility

The theory of possibility is not yet fully developed. The possibility measure is a special case of plausibility. Plausibility is a possibility measure if and only if the focal elements are nested /Du86/. Corresponding belief function is the necessity measure. In such a case the relation $Bel(A) \leq P(A) \leq Pl(A)$ can be rewritten as $N(A) \leq P(A) \leq \Pi(A)$.

If the degree of uncertainty for an event A is given in the form of the degree of possibility $\Pi_i(A)$, and the corresponding degree of necessity of $N_i(A)$, with reference to a source i (e.g. an opinion of the expert i), $\{\Pi_i(A), N_i(A)\}$, then the following formulas for combining degrees of uncertainty, $\{\Pi(A), N(A)\}$, are proposed by Dubois & Prade /Du88a/:

$$\Pi(A) = \frac{\min_i \Pi_i(A)}{\max \left\{ \min_i \Pi_i(A), \min_i (1 - N_i(A)) \right\}}$$

$$N(A) = 1 - \frac{\min_i (1 - N_i(A))}{\max \left\{ \min_i \Pi_i(A), \min_i (1 - N_i(A)) \right\}}$$

This rule of combination was proposed to extract the nonconflicting parts of the data from the different sources (e.g. given by experts) by following the conjunction operation defined in the fuzzy set theory (the denominator in the rules above carries out the renormalization). It was shown /Wu90/, however, that the combined measure of uncertainty can be sensitive to slight changes of individual opinions. It is because the rule of combination extracts only the nonconflicting part of opinions and renormalizes the resulting possibility and necessity function. When there is significant difference between experts, only a small portion of the opinions that remains nonconflicting is legitimate for use in combining information, and the renormalization of the resulting possibility and necessity functions greatly amplifies the slight changes in individual inputs.

There are also proposals to combine, using the theory of possibility, the possibility distributions given directly by experts /Du86/. Similarly to Dempster's rule of combination, the rules of combination used in the theory of possibility suffer from the lack of ability to allow for the evaluation of the credibility of the sources of information. There is no place in the rule where an analyst can assign measures of credibility to each individual.

Although some new ideas were proposed /Ga87b/ to overcome these difficulties by a generalization of the evidence theory on fuzzy sets, Dubois & Prade /Du88b/ have also attempted to apply alternate combination rules to cases where sources are unreliable, non-exhaustive, inconsistent, or dependent, proposed rules to combine information are still not mature for more complex real applications. In this respect the Bayesian approach is at present much more attractive in systems analysis /Wu90/.

4.4.3 Representing uncertainties and combining information in expert systems

The methods available for uncertainty representation and combining expert opinions described above are closely related to models of quantitative reasoning under uncertainty in knowledge-based expert systems /Ka86a, Le85, Le87, Le88, Ma85/.

Uncertainty arises constantly in practical situations and expert systems should be designed to solve real-world and engineering problems. Three basic types of situations, where inexact reasoning is necessary, are usually distinguished: first, "purely random" situations where a chance mechanism is directly involved; second, there are situations where possible randomness is due to incomplete information; and third, there are situations where uncertainty results from imprecisions of their characterization, because they cannot be described more precisely to eliminate those uncertainties.

In knowledge-based expert systems, uncertainty can manifest itself in data and in the inference process. Uncertainty in data usually results from the inaccuracies of empirical results, the subjectivity of certain judgements, and the vagueness of the concept used. In forming hypotheses, the rules of the form IF A THEN B or its variations are used and there are doubts about the reasoning process. So, to deal with uncertainty two fundamental issues need to be addressed: first, how degrees of uncertainty in factual data and inference rules should be represented, and second, how these degrees of uncertainty should be used in the inference process.

There are several known techniques in drawing inferences under uncertainty based on: Bayesian probability theory, Dempster-Shafer theory of evidence, Zadeh's fuzzy set theory and the related theory of possibility, and Mycin's model of confirmation and disconfirmations /Le85, Ma85, Ka86a, Le87, Du88a/. The advantages and drawbacks of the first three of these techniques will be discussed briefly, because they have the same background as those described above for combining expert opinions and are of interest for application within expert systems supporting PSA.

The Bayesian probability theory is well developed and contributes a firm basis for its application to modelling inexact reasoning. Bayesian theory has been successfully applied in many expert systems. However, there are significant drawbacks in applying this theory to expert systems /Le87/:

- (1) Owing to the subjective nature of probability assignment, there is no guarantee that a system of probabilities given by an expert is consistent and coherent. There are ways to eliminate such incoherence, but they can create a lot of overheads in terms of the expert's time and the computer's time.
- (2) The use of Bayes' theorem normally requires all the hypotheses to be disjoint. In a large expert system, dividing a solution space into mutually exclusive subsets may be impractical.

- (3) The assumption of a certain conditional independence in simplifying the formula for uncertain evidence may introduce errors into the system.
- (4) In the case of altering the probability of an event in a system (by adding or removing a hypothesis), we have to recalculate all the probabilities in order to preserve the coherence and consistency of the system.

The Dempster-Shafer theory is appealing because it is a natural mathematical generalization of the classical probability theory. It captures the fact that beliefs are generally not additive. However, the theory is fairly new and has not yet been widely applied. Therefore, the practical advantages of the theory are at present unclear. There are some problems in this theory which have not yet been resolved /Le87/:

- (1) A convincing interpretation of the theory has not yet been presented. The way the evidence is partitioned can affect the results in a nonintuitive manner. The meaning of Dempster's rule of combination is unclear except in very simple situations.
- (2) The numerical stability of the theory has not been analysed in great detail. In some cases, a small variation in the basic probability assignments can produce a large variation in the results.
- (3) Dempster's rule of combination cannot be applied in situations where there are considerable disagreements among the evidence, that is, when the cores of two belief functions are disjoint.
- (4) The theory, in its current form, may be inconvenient to use in the case of a long chain of inferences, i.e. when the conclusions of one inference become the premises of the next. As the number of steps in the inference chain increases, the recursive application of Dempster's rule becomes difficult because of the increasing complexity in the structure of the core of the resulting belief function.

The idea of fuzzy sets seems appealing because it allows imprecise linguistic terms to be represented and manipulated in a well-defined mathematical way. Traditional logic is extended to incorporate uncertainties by the introduction of fuzzy implication (IF A THEN B ELSE C) and fuzzy quantifiers (e.g. "some", "most"). There have recently been a number of expert systems and tools developed using fuzzy logic. There are, however, some serious drawbacks in applying fuzzy set theory to expert system problem solving /Le87/:

- (1) Membership functions are context-sensitive. Blind applications of general, generic membership functions to all situations is undesirable.
- (2) Depending on the structure of individual membership functions, the computational complexity of a fuzzy logical system can vary significantly.
- (3) Owing to the lack of canonical examples to provide interpretations for fuzzy results, it is difficult to interpret a membership function obtained from fuzzy logical implication when the resulting function does not resemble any predefined membership functions the system starts with.

Thus, there is no best technique of uncertainty representation and reasoning in expert systems. The choice of the technique depends on the problem to be solved in a given domain /Le87/.

4.4.4 Treatment of uncertainties in computerized PSA/HRA

The PSA methodologies currently applied in engineering practice are based on the Bayesian probabilistic framework. As was mentioned in some new PSA studies, e.g. NUREG-1150, extensive engineering judgement has been applied and the results obtained depend, therefore, heavily on expert opinions. The practice of producing data by experts for probabilistic analyses

has been criticized lately /Mo88/. This concerns in particular the engineering judgement involved in modelling human reliability /Do90/.

Only seldom can the probabilistic data be referred to as "hard" data, based on "solid" statistical data. Usually in PSA practice the reliability data are rather "soft", picked up by experts from various data bases from records not for identical but similar components. The acquisition and evaluation of these reliability data also depend on engineering judgement. It concerns in particular quantification of human reliability, because the wide spectrum of possible human behaviour and the error mechanisms potentially involved make it too costly or even impossible to base the analyses on the extensive use of experimental statistical data. Thus, various human reliability modelling techniques described in Item 2.3 including psychological scaling methods are usually applied. They require more or less extensive expert judgement depending on the case considered and model selected.

In such a situation the application of the Bayesian framework for complex probabilistic assessments is associated with very strong assumptions (see remarks above concerning the drawbacks of the Bayesian framework for designing an expert system). As was mentioned, uncertainties associated with probabilistic modelling are not only of a random nature but often associated with limited knowledge or imprecise statements involved. There are also assumptions made in several steps of the PSA process, e.g. concerning human reliability modelling /IA92b/, which can significantly influence the probabilistic results obtained. In such a situation we think that considering an alternative, less restrictive, framework to perform the probabilistic assessments, is, at least in some practical situations, justified.

We propose to develop software for human reliability modelling and probabilistic modelling of the plant based on the two following approaches to represent uncertainty and combining information:

- (1) Conventional Bayesian framework.
- (2) Approximate combined framework with fuzzy probabilistic results.

The conventional Bayesian, more restrictive, framework will be used when "hard" reliability data for components are available or extensive multi-expert techniques are applied in cases of rare hardware failure events and human errors not supported by statistical data. This approach will be applied for conventional logical modelling of the plant for initiating events considered.

In the approximated combined framework we propose to take advantage of the approaches described above to deal with uncertainty representation and combining of information as follows:

- (a) to combine conflicting or non-equivalent information using the Bayesian approach;
- (b) having the probabilistic distributions or parameters of distributions (e.g. mean or standard deviation) for failure events to evaluate their possibilistic distributions, i.e. fuzzy probability measures represented by fuzzy numbers, e.g. in the triangular form /Te83, Ka86b/;
- (c) using fuzzy numbers calculus /Ka86b, Ka86c/ to quantify the fault and event trees based on relevant minimal cut sets;
- (d) to analyse the results obtained and to prepare information for decision making or further analyses based on fuzzy measures of uncertainties.

This approximate framework would be consistent to include the future additional features, e.g. to combine non-conflicting information represented as Shafer's belief structures using Dempster's rule of combination /De67/ or to combine non-conflicting information represented

by the possibilistic measures of probability using the relevant rule of combination. There are expectations that new proposals may emerge from the the non-probabilistic theories described above which will be interesting to apply in PSA practice /Wu90/.

The structure of logical modes (the fault and event trees) would be developed in a conventional way based on "hard" topological and functional-logical information. The next stage of the system development would be to also consider uncertainty in the deterministic object-oriented modelling and the classification of plant's states (with the analysis of their influence on logical models of the plant and the assessment of consequences). The whole process of the logical modelling of the plant would be supported by an expert system with some functions automated, (e.g. the construction of the fault/event trees on the knowledge acquired from the domain experts).

Opinions have been expressed that the fuzzy set theory is well suited to model man-machine systems. Approaches based on the fuzzy set theory are used in the diagnostic software for complex systems /Gu87, Gu88, Gu89/. This theory is applied for human problem solving /Hu84, Sh88/ and advanced modelling of operator behaviour /Am87/. There are also proposals to model human reliability using fuzzy probabilistic measures /Te83/. Some ideas to apply the combined framework for the fuzzy probabilistic quantification within PSA have been proposed in ref. /Ko90/. The fuzzy probabilistic approach is well suited for quantification of such logical structures as HRAET and OAT.

The simplest way to convert the a quasi-random variable Q_i , characterized by the pdf $\delta_i(q)$, into a fuzzy variable $\tilde{Q}_i$ is to divide $\delta_i(q)$ by $\max_q \delta_i(q)$ /Ka86b/:

$$\mu_i(q) = \frac{\delta_i(q)}{\max_q \delta_i(q)}$$

However, a simplified shape of the fuzzy variable as a triangular fuzzy number can be proposed, described by a set of three real numbers (q_l^i, q_m^i, q_n^i) (according to $\delta_i(q)$) with the following assignments:

- $q_l^i = q_{5\%}^i$ (5% percentile of i-th quasi-random variable);
- $q_m^i = q_{50\%}^i$ (50% percentile, i.e. the median value);
- $q_n^i = q_{95\%}^i$ (95% percentile).

These real numbers correspond to the lower bound, median and higher bound of the probability, and in some cases (e.g. for log-normal distributions) the calculation of the error factor EF_i , defined as $EF_i = q_m^i / q_l^i = q_n^i / q_m^i$ is useful. Other assignments can be proposed, e.g. based on a mean value and standard deviation of a quasi-random variable. Thus, such fuzzy numbers can be represented in some cases using two real numbers: (q_i, EF_i) . The possibilistic measures in the form of the fuzzy probability can sometimes be obtained directly by combining multi-expert judgements /Du86/ and then represented by the fuzzy number (q_l^i, q_m^i, q_n^i) .

Fuzzy quantification of a fault tree (or an event tree) is very simple. Having the set (q_l^i, q_m^i, q_n^i) for each fuzzy variable $\tilde{Q}_i$, the fuzzy probability of the top event is calculated according to the general formula:

$$\tilde{Q} = 1 \ominus \prod_{j=1}^m (1 \ominus \prod_{i \in M_j} \tilde{Q}_i)$$

where: m is a number of minimal cut sets, M_j is j -th minimal cut set; $\ominus$ and Π are special operators on fuzzy numbers; for subtraction and multiplication respectively

Multiplying fuzzy numbers in the form proposed and their subtraction from 1 is very simple and can be performed in a similar way to that described in ref. /Ta83/. The calculation cost of the fuzzy probability $\tilde{Q}$ is much lower than in the case of quasi-random variables, although the uncertainty range obtained will be wider. For example, if two fuzzy numbers representing the fuzzy probabilities: (q_1, EF_1) and (q_2, EF_2) are multiplied, then a combined fuzzy number is $(q_1 q_2, EF_1 EF_2)$. Thus, the error factor of the new fuzzy probability is a product of the error factors of these two fuzzy numbers.

The approach proposed seems to be fully justified when data for calculation are weakly supported by statistics and are based mainly on expert opinions. This is precisely the case when assessment of the human factor is involved. Some proposals to use the fuzzy PSA results for safety-related decision making are described in ref. /Ko90/.

Although some authors are sceptical as regards the application of new non-Bayesian theories in PSA practice /Wu90/, they do not exclude this in the future, when these new theories will be more mature. On the other hand the fuzzy measures have already been proposed in PSA practice as the design thresholds limits to be compared with quantitative results /Ap89, Ko90/. The proposal described above can be treated as a small step in the practical application of a non-Bayesian theory. This direction is also promising in building the PSA expert systems when the uniform treatment of uncertainties in data and knowledge will be of interest.

Application of the fuzzy set theory approach is often considered to be attractive and effective for decision making, and is justified especially on the higher level of risk related decision making /Vo86/.

4.5 Proposals for system development

The research and software design effort should be aimed at developing a software environment (Fig. 4.3) supporting the PSA/HRA studies of scope I and II (Table 4.1). This requires the integration of several programming applications to be developed based on:

- (1) Hybrid expert system shell (and AI supporting software);
- (2) CAD (Computer Aided Design) to support the graphical representing of topological and functional-logical information;
- (3) Data base (e.g. Reliability Data Base) and its management compatible with applications applied;
- (4) Model base and its management;
- (5) Interface to the user.

Other works mentioned earlier to be carried out is as follows:

- the development of an HRA expert system using conventional HRA techniques (THERP and other methods as proposed in Item 3.6),
- knowledge acquisition for the PSA expert system to enable automatic construction of the fault trees (e.g. by means of backward chaining) and event trees (e.g. by means of forward chaining) based on topological/functional-logical knowledge base (Fig. 4.3),
- integration of the HRA expert system with PSA expert system,
- the development of a framework for representing and treating uncertain data/knowledge including expert opinions,
- verification and validation of the entire HRA/PSA expert system.

The scope of work described is very complex and requires considerable research and designing effort to be spread over a longer period /IA90, IA91a/. One of the most important aspects in building the knowledge-based systems (expert systems) is their verification and validation /IA91b/ (see also Appendix 1).

5 Conclusions

Human Reliability Analysis (HRA) is considered as one of the most important tasks within Probabilistic Safety Analysis (PSA), because the so-called human factor has been recognized as the main contributor to the risk associated with operation of complex technological hazardous systems. Many models for quantifying human reliability have been developed to date, but only a few, for example the Technique for Human Error Rate Prediction (THERP), are more widely used in PSA studies of nuclear power plants.

A methodological difficulty of human error analysis is inherited from the special nature of mind. Human beings are distinct from the technical components because their behaviour is characterized by substantially larger variability and complexity. Therefore in the development of the operator behaviour models one has to include technical as well as psychological aspects. Contemporary human error analysis suffers from two main unresolved issues: the sparse data base on all human activities and the shortage of methods for dealing with the intention behind 'extraneous' actions.

In addition to these 'theoretical' aspects there exists a substantial problem which has an influence on HRA quality. Insights from comparisons of the HRA model applications show that different analysis teams produce different results for the same problem /Po88a/. The quantifying uncertainties in human reliability can be reduced by using the supporting computer-based tools.

Thus a software system supporting HRA is needed. The system should be designed on the basis of expert system technology and developed with regard to psychological cognitive theories to cover important cases of human interactions/errors. In the report an approach is proposed to develop the software system in consecutive steps based upon several modelling techniques for three levels of the HRA/PSA effort. The interface of this system should be designed as "user friendly" to help the analyst in selecting the relevant modelling techniques for different kinds of man-machine interactions including important types and effects of actions/errors.

After the computer implementation of the THERP technique it is proposed to develop the computerized versions of following techniques: ASEP-HRAP, INTENT, SLIM and APJ as well as an integrating module using the expert system technology. The software system should be developed in such a way as to enable communication with various data bases. The evaluation of HEPs for given attributes/parameters of the model selected should be documented in Human Reliability Data Base (HRDB). Thus, the data obtained can be scrutinized and easily modified when new evidence is available. The system might be used for documenting of HRA results for both the single- and multi-expert HRA techniques. The HRA expert system proposed could in the future be an important part of the PSA expert system.

The PSA expert system (knowledge-based system) should consist of several modules which enable full documenting of the analyses with some functions to be performed automatically. There are several related research topics to be undertaken. The scope of research and designing works is very complex, therefore they have to be spread over a longer time with a contribution by several domain experts and software designers.

REFERENCES

- /Am87/ Amendola, A., Bersini, U., Cacciabue, P.C., Mancini, G. (1987). Modelling operators in accident conditions: advances and perspectives on a cognitive model. *International Journal on Man-Machine Studies*, Vol. 27, pp. 599-612.
- /Ap88b/ Apostolakis, G.E., Bier, V.M., Mosleh, A. (1988). A critique of recent models for human error rate assessment. *Reliability Engineering and System Safety*, Vol. 22, pp. 201-217.
- /Ap89/ Apostolakis, G.E. (1989). Uncertainty in probabilistic safety assessment. *Nuclear Engineering and Design*, Vol. 115, pp. 173-179.
- /Be80/ Bello, G.C., Colombardi, V. (1980). The human factors in risk analyses of process plants: The control room operator model "TESEO". *Reliability Engineering*, Vol.1, pp.3-14.
- /Be87/ Bezdek, J.C. (1987). *Pattern Recognition with Fuzzy Objective Function Algorithms*. Plenum Press, New York and London.
- /Be88/ Bersini U., Cacciabue P.C., Mancini G. (1988). Cognitive modelling: a basic complement of human reliability analysis. *Reliability Engineering and System Safety*, Vol. 22, pp. 107-128.
- /Bi92/ Bieniarz, P.B. (1992). PCPPS (PC Parallel Processing Simulation). Risk Management Associates, Inc., Albuquerque, New Mexico 87107.
- /Ca88/ Cacciabue, P.C. (1988). Evaluation of human factors and man-machine problems in the safety of nuclear power plants. *Nuclear Engineering and Design*, Vol. 109, pp. 417-431.
- /Ca89/ Cacciabue, P.C., Decortis, F., Mancini, G., Masson, M., Nordvik, J.P. (1989). A cognitive model in a blackboard architecture: synergism of AI and psychology. *Proceedings of the Second European Meeting on Cognitive Science Approaches to Process Control*, Sienna, Italy, October 24-27, 1989. Reprint in *Reliability Engineering and System Safety*, Vol.36, 1992, pp. 187-197.
- /Da79/ Davis R. (1979). Interactive transfer of expertise: acquisition of new inference rules. *Artificial Intelligence*, Vol. 12, pp. 121-157.
- /Da80/ Davis R. (1980). Meta-rules: reasoning about control. *Artificial Intelligence*, Vol. 15, pp. 179-222.
- /De67/ Dempster, A.P. (1967). Upper and lower probabilities induced by a multivalued mapping. *Annual Mathematical Statistics*, Vol. 38, pp.325-339.
- /De68/ Dempster, A.P. (1968). A generalization of Bayesian inference. *Journal of the Royal Statistical Society, Series B*, Vol. 30, pp. 205-247.
- /Do88/ Dougherty, E.M., Fragola J.R. (1988). *Human Reliability Analysis: A Systems Engineering Approach with Nuclear Power Plant Applications*. A Wiley-Interscience Publication, John Wiley & Sons Inc., New York.
- /Do90/ Dougherty, E.M. (1990). Human reliability analysis - where should thou turn ? (Guest Editorial). *Reliability Engineering and System Safety*, Vol. 29.
- /Du86/ Dubois, D., Prade, H. (1986). Fuzzy sets and statistical data. *European Journal of Operational Research*, Vol. 25, pp. 345-356.
- /Du88a/ Dubois, D., Prade, H. (1988). *Possibility Theory: An Approach to Computerised Processing of Uncertainty*. Plenum Press, New York.
- /Du88b/ Dubois, D., Prade, H. (1988). Representation and combination of uncertainty with belief functions and possibility measures. *Computational Intelligence*, Vol. 4, pp. 244-64.

/Ed68/ Edwards, W. (1968): Information Processing, Decision Making and Highway Safety, in: J. O'Day (ed.): Driver behavior: Cause and Effect. Washington, D.C.: Insurance Institute for Highway Safety

/Em83/ 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.

/Em84/ Embrey, D.E., Humphreys, P., Rosa, E.A., Kirwan, B., Rea, K. (1984). SLIM-MAUD: An approach to assessing human error probabilities using structured expert judgment. NUREG/CR-3518, March 1984.

/Em86/ Embrey, D.E. (1986). SHERPA: A systematic human error reduction and prediction approach. Proceedings of the International Topical Meeting on Advances in Human Factors in Nuclear Power Systems, American Nuclear Society, April 21, 1986.

/FI75/ Fleming, K.N. et al. (1975). HTGR Accident Initiation and Progression Analysis Status Report, Vol. II, AIPA Risk Assessment Methodology. GA/A13617, San Diego, CA.

/Fu92/ Fujita, Y.(1992). Human reliability analysis: A human point of view. Reliability Engineering and System Safety, Vol.38, pp. 71-79.

/Ga76/ Gaines, B.R. (1976). Foundations of fuzzy reasoning. International Journal of Man-Machine Studies, Vol. 8, pp. 623-68.

/Ga87a/ Garriba, S.F., Guagnini, E., Mussio, P. (1987). Knowledge paradigm and error management in an expert system for fault-tree construction. PSA'87 International Topical Conference on Probabilistic Safety Assessment and Risk Management, Zurich, Switzerland, Aug. 30 - Sept. 4, 1987.

/Ga87b/ Garriba, S.F., Servida, A., Apostolakis, G. (1987). Fuzzy models for evaluating expert judgment. PSA'87 International Topical Conference on Probabilistic Safety Assessment and Risk Management, Zurich, Switzerland, Aug. 30 - Sept. 4, 1987.

/Ge89/ Gertman, D.I., Gilmore, W.E., Galyean, W.J., Groh, M.R., Gentillon, C.D., Gilbert, B.G. (1989). Nuclear Computerised Library for Assessing Reactor Reliability, Vol. I: Summary description and Vol. V: The Data Manual. (with W.J. Reece & G.B. Beers) NUREG/CR-4639, US NRC Washington, DC, February 1989.

/Ge92/ Gertman, D.I., Blackman, H.S., Haney, L.N., Seidler, K.S., Heidi, A.H. (1992). INTENT: a method for estimating human error probabilities for decision-based errors. Reliability Engineering and System Safety, Vol.35, pp.127-136.

/Gu87/ Guth, M.A.S. (1987). Incorporating "Fuzzy" Data and Logical Relations in the Design of Expert Systems for Nuclear Reactors. Oak Ridge National Laboratory, Conf-870832-3, 014410.

/Gu88/ Guth, M.A.S. (1988). Uncertainty analysis of rule-based expert systems with Dempster-Shafer mass assignments. International Journal of Intelligent Systems, Vol. 3, pp. 123-139.

/Gu89/ Guth, M.A.S. (1989). Some uses and limitations of fuzzy logic in artificial intelligence reasoning for reactor control. Nuclear Engineering and Design, Vol. 113, pp. 99-109.

/Ha82/ Hall, R.E., Fragola, J., Wreathall, J. (1982). Post Event Human Decision Errors: Operator Action Tree / Time Reliability Correlation. NUREG/CR-3010, Washington, D.C.

/Ha83/ Hayes-Roth, F., Waterman, D.A., Lenat, D.B. (eds.) (1983). Building Expert Systems. Addison-Wesley Publishing Company, Inc.

/Ha84/ Hannaman, G.W., Spurgin, A.J. (1984). Systematic Human Action Reliability Procedures (SHARP). EPRI NP-3583, Research Project 2170-3, Palo Alto.

- /Ha85a/ Hayes-Roth B. (1985). A blackboard architecture for control. *Artificial Intelligence*, Vol. 26, pp. 251-321.
- /Ha85b/ Hannaman, G.W., Spurgin, A.J., Lukic, Y. (1985). A model for assessing human cognitive reliability in PRA studies. *IEEE Third Conference on Human Factors and Nuclear Power Plants*, Monterey, California, June 23-27.
- /Ha88/ Hannaman, G.W., Worledge, D.H. (1988). Some developments in human reliability analysis approaches and tools. *Reliability Engineering and System Safety*, Vol. 22, pp. 235-256.
- /Ho88/ Hoyos, C. Graf / Zimolong, B. (1988): "Occupational Safety and Accident Prevention: Behavioural Strategies and Methods". Elsevier Amsterdam et al.; (Vol.11 of "Advances in Human Factors/Ergonomics").
- /Ho89/ Holmberg J., Silvennoinen P., Vira J. (1989). Application of the Dempster-Shafer theory of evidence for accident probability estimates. *Reliability Engineering and System Safety*, Vol. 26, pp. 47-58.
- /Hu84/ Hunt R.M., Rouse W.B. (1984). A fuzzy rule-based model of human problem solving. *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. SMC-14, No.1, pp. 112-120.
- /Hu88/ Humphreys, P. (Ed.) (1988). *Human Reliability Assessors Guide*. Safety and Reliability Directorate, RTS 88/95Q, October 1988.
- /IA90/ IAEA (1990). IAEA-TECDOC-542, Use of Expert Systems in Nuclear Safety. Report of a Technical Committee Meeting organized by the International Atomic Energy Agency held in Vienna, 17-21 October 1988.
- /IA91a/ IAEA (1991a). IAEA-J4-TC-660.2, Practical Experience with Expert System Applications in Nuclear Safety. Report of a Technical Committee Meeting/Workshop organized by the International Atomic Energy Agency held in Moscow, 1-5 October 1990.
- /IA91b/ IAEA (1991b). IAEA-J4-RC-471, Development of Safety Related Expert Systems. Report of the First Research Co-ordination Meeting of the Co-ordinated Research Programme organized by the International Atomic Energy Agency and held in Vienna, 4-8 March 1991.
- /IA91c/ IAEA-TECDOC-592 (1991). Case study on the use of PSA methods: Human reliability analysis. IAEA, Vienna, April 1991.
- /IA92a/ IAEA (1992a). Safety Series No. 50-P-4. Procedures for Conducting Probabilistic Safety Assessment of Nuclear Power Plants (Level 1).
- /IA92b/ IAEA (1992b). Procedure for Conducting Human Reliability Analysis in Probabilistic Safety Assessment (Draft of a report).
- /Iv92/ Iven, F.W., Mertens, J. (1992). State of the art of level-1 PSA methodology. Chapter 2: Human Factors. Final draft of a report prepared for CSNI/PWG5, Organization for Economic Co-operation and Development, Nuclear Energy Agency.
- /Ka72/ Kahneman, D., Tversky, A. (1972). Subjective probability: a judgment of representativeness. *Cognitive Psychology*, Vol. 3, pp. 430-454.
- /Ka86a/ Kanal, L.N., Lemmer, J.F. (Eds.) (1986). *Uncertainty in Artificial Intelligence*. Elsevier Science Publishers B.V., North-Holland, Amsterdam.
- /Ka86b/ Kandel, A. (1986). *Fuzzy Mathematical Techniques with Applications*. Addison-Wesley Publishing Company, Massachusetts.
- /Ka86c/ Kaufmann, A. (1986). Hybrid data - various associations between fuzzy subsets and random variables, pp. 171-211. In: "Fuzzy Sets Theory and Applications", A. Jones, et al. (eds.), D. Reidel Publishing Company.

- /Ka87/ Karvonen, I., Suokas, J. (1987). HAZOPEX - expert system supporting the safety analysis. A paper presented at the Society of Reliability Engineering Symposium (Scandinavian Chapter), Helsingor, Denmark.
- /KA92/ KAPPA-PC Reference Manual (1992). KAPPA-PC Version 2.0. Intellicorp, Inc.
- /KE90/ KEE (Knowledge Engineering Environment)(1990). Software Development System, Version 3.1, User's Manual (Vol. 1,2,3). Intellicorp., Mountain View, CA, USA.
- /Ko90/ Kosmowski, K.T. (1990). Expert opinions in PSA and decision making under uncertainties. A paper presented during 7th International Conference on Reliability and Maintainability, Brest, France, June 1990.
- /Ko92/ Kosmowski, K.T., Duzinkiewicz, K. (1992). An integrated approach in probabilistic modelling of hazardous technological systems with emphasis on human factor. Manuscript prepared for ICOSAR'93, 6th International Conference on Structural Safety and Reliability (Innsbruck, Austria, 9-13 August 1993).
- /Le85/ Lesmo L., Saitta L., Torasso P. (1985). Evidence combination in expert systems. International Journal of Man-Machine Studies, Vol. 22, pp. 307-326.
- /Le86/ Leao L.V., Talukdar S.N. (1986). An environment for rule-based blackboards and distributed problem solving. Artificial Intelligence, Vol. 1, pp. 70-79.
- /Le87/ Lee, N., Grize, Y.L., Dehnad, K. (1987). Quantitative models for reasoning under uncertainty in knowledge-based expert systems. International Journal of Intelligent Systems, Vol. II, pp. 15-38.
- /Le88/ Lemmer, J.F., Kanal, L.N. (Eds.)(1988). Uncertainty in Artificial Intelligence 2. North-Holland, Amsterdam.
- /Ma85/ Martin-Clouaire R., Prade H. (1985). On the problems of representation and propagation of uncertainty in expert systems. International Journal of Man-Machine Studies, Vol. 22, pp. 251-264.
- /Ma90/ Magne, L. (1990). A workstation for reliability engineers. Working Material on Practical Experiences with Expert System Applications in Nuclear Safety, Report of a Technical Committee Meeting/Workshop organized by the IAEA, held in Moscow, 1-5 October 1990, IAEA-J4-TC-660.2, Reproduced by the IAEA, Vienna, Austria, 1991.
- /Me89/ Melis, M., Barbet, J.F., Olivieri, M. (1989). Computerised safety and reliability assessment: from computer codes development to expert system techniques application. Reliability Engineering and System Safety, Vol. 25, pp. 175-181.
- /Me90/ Mennell, S.J., Raymond, D.P., Reeves, A. B. (1990). Probabilistic Safety Assessment (PSA) applications using knowledge based systems techniques. Report of a Technical Committee Meeting/Workshop organized by the IAEA, held in Moscow, 1-5 October 1990, IAEA-J4-TC-660.2, Reproduced by the IAEA, Vienna, Austria, 1991.
- /Mo77/ Morris P.A. (1977). Combining expert judgments: a Bayesian approach. Management Science, Vol. 23, pp. 679-693.
- /Mo88/ Mosleh, A., Bier, V.M., Apostolakis, G. (1988). A critique of current practice for the use of expert opinions in probabilistic risk assessment. Reliability Engineering and System Safety, Vol. 20, pp.63-85.
- /Na89/ Naser, J.A., et al. (1989). Expert system applications to nuclear power for enhancement of productivity and performance. Nuclear Engineering and Design, Vol. 117, pp. 275-286.
- /Na90/ Naser, J.A. (Ed.) (1991). Expert Systems Applications for the Electric Power Industry. Vol. 1 and 2, Electric Power Research Institute, Hemisphere Publishing Corporation, New York.

- /Ne87/ Neuschaefer, C.H. (1987). Application of C-E'S generic diagnostic system to power plant diagnostics. Proceedings: 1987 Conference on Expert-System Applications in Power Plants, EPRI CS-6080, Project 2923-1, December 1988.
- /Ni86/ Nii H.P. (1986). Blackboard systems: the blackboard model of problem solving and the evolution of blackboard architectures. The AI Magazine (Summer 1986), pp. 38-53.
- /NU85/ NUREG-1154 (1985): Loss of Main and Auxiliary Feedwater Event at the Davis-Besse Plant on June 9, 1985. Washington, D.C.
- /NU87/ NUREG-1150 (1987). Reactor risk reference document, Washington, D.C.
- /Ph89/ Philips, L.D., Humphreys, P., Embrey, D., Selby, D.L. (1989). A Socio-Technical Approach to Assessing Human Reliability). Decision Analysis Unit, The London School of Economics and Political Science, University of London.
- /Po81/ Potash, L. M., Steward, M., Dietz, P. E., Lewis, C. M., Dougherty, E. M. jr (1981). Experience in integrating the operator contribution in the PRA of actual operating plants. Proceedings of the International Topical Meeting on Probabilistic Risk Assessment, Port Chester, N.Y.
- /Po82/ Poulton, E.C. (1982). Biases in quantitative judgments. Applied Ergonomics, Vol. 13.1, pp. 31-42.
- /Po88/ Poucet, A. (1988): "Survey of Methods Used to Assess Human Reliability in Human Factors Reliability Benchmark Exercise", Reliability Engineering & System Safety 22, p.257-268.
- /Po88a/ Poucet, A. (1988). Human Factors Reliability Benchmark Exercise (HF-RBE). Final Report - part I, Summary of Results and Conclusions. P.E.R.1482/88, Commission of the European Communities, Joint Research Centre Ispra, Ispra (Va), Italy.
- /Po88b/ Poucet, A. (1988). Survey of methods used to assess human reliability in human factors reliability benchmark exercise. Reliability Engineering and System Safety, Vol. 22, pp. 257-268.
- /Po90/ Poucet, A. (1990). STARS: knowledge based tools for safety and reliability analysis. Reliability Engineering and System Safety, Vol. 30, pp. 379-397.
- /Ra83/ Rasmussen, J. (1983). Skills, rules, knowledge; signals, signs and symbols and other distinctions on human performance models. IEEE Transaction on Systems, Man and Cybernetics, SMC - 13, No. 3.
- /Ra87/ Rasmussen, J., Duncan, K., Leplat, J. (Eds.)(1987). New Technology and Human Errors, John Wiley & Sons, London.
- /Re87/ Reason, J. (1987). 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. Leaflet (Eds), John Wiley and Sons, Chichester (UK) 1987.
- /Re90a/ Reason, J. (1990): "Human Error", Cambridge et al.: Cambridge University Press.
- /Re90b/ Reeves, A.B., Mennell, S.J. (1990). A review of expert system development tools. Report of a Technical Committee Meeting/Workshop organized by the IAEA, held in Moscow, 1-5 October 1990, IAEA-J4-TC-660.2, Reproduced by the IAEA, Vienna, Austria, 1991.
- /Re92a/ Reer B. (1992). A probabilistic method proving the reliability effect of time and organizational factors. Proceedings of the Eighth EURO Summer Institute on Risk Management in Complex Production and Transportation Systems, Rosersberg Castle near Stockholm, Ed.: Swedish OR Association, c/o FOA, S17290 Sundbyberg
- /Re92b/ Reer B. (1993). Entscheidungsfehler des Betriebspersonals von Kernkraftwerken als Objekt probabilistischer Risikoanalysen. Doctoral thesis, D468, BUGH Wuppertal

- /Re93/ Reer, B., Degen, G., Mertens, J. (1993) ESAP, ein benutzerfreundliches Expertensystem zur systematischen Analyse von Personalhandlungen in Risikostudien. Tagungsbericht Jahrestagung Kerntechnik 93, inforum Verlagsgesellschaft, Bonn
- /Ro88/ Rolston D.W. (1988). Principle of Artificial Intelligence and Expert Systems Development. McGraw-Hill Book Company.
- /Sa87/ Salvendy, S. (1987). Handbook of Human Factors, A Wiley-Interscience Publication. John Wiley & Sons, New York.
- /SAS89/ SAS Institute Inc.: SAS® Languages and Procedures: Usage, Version 6, First Edition. Cary, N.Y., 1989
- /Se83/ Seavers, D.A., Stillwell, W.G. (1983). Procedures for Using Expert Judgment to Estimate Human Error Probabilities in Nuclear Power Plants Operations. NUREG/CR-2743, US NRC.
- /Sh76/ Shafer, G. (1976). A Mathematical Theory of Evidence. Princeton University Press.
- /Sh88/ Sheridan, T.B., Ren, J. (1988). Application of fuzzy set theory in simulation of human operator safety monitoring. IFAC Man-Machine Systems, Oulu, Finland, pp. 421-426.
- /Si82/ Sixtl, F. (1982). Messmethoden der Psychologie. Theoretische Grundlagen und Probleme, Beltz Verlag.
- /Si84/ Siegel, A., et al. (1984). Maintenance Personnel Performance Simulation (MAPPS) Model. NUREG/CR-3626, Vol. 1 and 2, US NRC.
- /Sw80/ Swain, A.D. / Guttmann, H.E. (1980): Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Applications. Draft Report, NUREG/CR-1278, Washington, D.C.
- /Sw83/ Swain, A.D., Guttmann, H.E. (1983). Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Application. Final Report, NUREG/CR-1278.
- /Sw87/ Swain, A.D. (1987). Accident Sequence Evaluation Procedure, Human Reliability Analysis Procedure. Sandia National Laboratories, NUREG/CR-4772 (SAND86-1996), U.S. Nuclear Regulatory Commission, Washington, DC, February 1987.
- /Sw88/ Swain, A.D. (1988). Human Reliability Analysis Methods. International Seminar on Probabilistic Methods in Reliability and Risk Evaluation, Dubrovnik - Cavtat, Yugoslavia, September 12-16, 1988.
- /Ta83/ Tanaka H., et al. (1983). Fault-tree analysis by fuzzy probability. IEEE Transactions on Reliability, Vol. R-32, No.5, pp. 453-457.
- /Te83/ Terano T., Murayama Y., Akiyama N. (1983). Human reliability and safety evaluation of man-machine systems. Automatica, Vol.19, No.6, pp.719-722.
- /Tu90/ Turban E. (1990). Decision Support and Expert Systems: Management Support Systems. Macmillan Publishing Company, New York.
- /Tv73/ Tversky, A., Kahneman, D. (1973). Availability: a heuristic for judging frequency and probability. Cognitive Psychology, Vol. 5, pp. 207-232.
- /Vo86/ Volta, G., Orway, H. (1986). The logic of Probabilistic Risk Assessment versus decision levels. Nuclear Engineering and Design, Vol. 93, pp.329-334.

- /Wa87/ Wakefield,D.J. / Adams,C.D. (1987): Quantification of Dynamic Errors in the TMI-1 PRA, in: ENS/SNS (Ed.): "Probabilistic Safety Assessment and Risk Management. PSA'87", Verlag TUV Rheinland Koeln; 1987.
- /Wa88a/ Wakefield, D.J. (1988). Application of the human cognitive reliability model and Confusion Matrix approach in a Probabilistic Risk Assessment. Reliability Engineering and System Safety, Vol. 22, pp. 295-312.
- /Wa88b/ Wang, J., Modarres, M., Hunt, R.N.M. (1988). Probabilistic Risk Assessment: a look at the role of Artificial Intelligence. Nuclear Engineering and Design, Vol. 106, pp. 375-387.
- /Wa90/ Wang, J., Modarres, M. (1990). REX: An intelligent decision and analysis aid for reliability and risk studies. Reliability Engineering and System Safety, Vol. 30, pp. 195-218.
- /Wi81/ Winkler R.L. (1981). Combining probability distributions from dependent information sources. Management Science, Vol. 27, pp. 479-488.
- /Wi85/ Williams, J.C. (1985). HEART - A proposed method for achieving high reliability in process operation by means of human factors engineering technology. Proceedings of a Symposium on the Achievement of Reliability in Operating Plant, Safety and Reliability Society, 16 September 1985.
- /Wi88/ Williams, J.C. (1988). A data-based method for assessing and reducing human error to improve operational experience. Proceedings of IEEE 4th Conference on Human Factors in Power Plants, Monterey, California, 6-9 June 1988.
- /Wo86/ Woods, D.D., et al. (1986). Models of Cognitive Behaviour in Nuclear Power Plant Personnel. NUREG/CR-453, Vol. 1 and 2.
- /Wo87/ Woods, D.D., et al. (1987). Cognitive Environment Simulation: an Artificial Intelligence System for Human Performance Assessment. NUREG/CR-486, Vol. 1, 2 and 3.
- /Wo87a/ Woods,D.D. / Roth,E. / Pople,H. jr. (1987). Cognitive Environment Simulations: An Artificial Intelligence System for Human Performance Assessment. Vol.2: Modelling Human Intention Formation, NUREG/CR-4862, Washington, D.C.
- /Wo88/ Woods, D.D., Roth, E.M. (1988). Modelling human intention formation for human reliability assessment. Reliability Engineering & System Safety, Vol. 22, pp. 169-200.
- /Wo90b/ Woods,D.D. / Roth,E. / Pople,H. jr. (1990): "The Cognitive Environment Simulation as Tool for Modelling Human Performance", Main Report, NUREG/CR-5213, Vol.2, Washington, D.C.
- /Wr82/ Wreathall, J.W. (1982). Operator Action Tree, An Approach to Quantifying Operator Error Probability During Accident Sequences. NUS Report 4159, NUS Corporation, Gaithersberg, Maryland, July 1982.
- /Wu90/ Wu, J.S., Apostolakis, G.E., Okrent, D. (1990). Uncertainties in system analysis: probabilistic versus nonprobabilistic theories. Reliability Engineering and System Safety, Vol. 30, pp. 163-181.
- /Za78/ Zadeh, L.A. (1978). Fuzzy sets as a basis for theory of possibility. Fuzzy Sets and Systems, Vol. 1, pp. 2-28.
- /Za83/ Zadeh L.A. (1983). The role of fuzzy logic in the management of uncertainty in expert systems. Fuzzy Sets and Systems, Vol. 11, pp. 199-227.
- /Za86/ Zadeh L.A. (1986). A simple view of the Dempster-Shafer theory of evidence and its implication for the rule of combination. The AI Magazine (Summer 1986), pp. 85-90.

Appendix 1

Features of the expert system technology

A1.1 Expert Systems and Decision Support Systems

Expert systems (ESs) are computerized advisory programmes that attempt to imitate or substitute the reasoning process and knowledge of experts in solving specific types of problems /Tu90/.

The characteristics and capabilities of ESs make them different from conventional computer systems. For example, their knowledge base is separated from the processing (inference) mechanism, execution is done by using heuristics and logic, explanation is a part of most ESs, they can operate with incomplete or uncertain information and can deal with qualitative data, etc.

Expert systems are composed of two major parts: the development environment and the consultation (run-time) environment (Fig. A1.1) The development environment is used by the ES builder to introduce expert knowledge into the knowledge base. The consultation environment is used by a user, who may be a nonexpert, to obtain expert knowledge and advice.

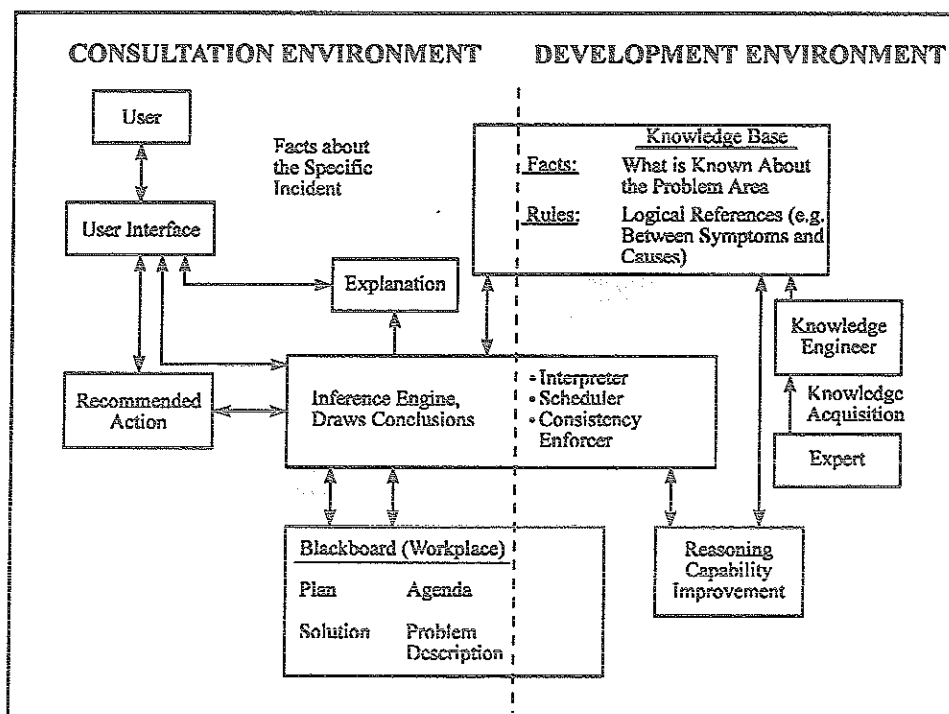


Fig. A1.1: Structure of an expert system /Tu90/

The following components exist in a more sophisticated expert system: knowledge acquisition (acquired by a knowledge engineer from an expert), knowledge base, inference engine, blackboard (workplace), user interface, explanation facility, and knowledge refining system /Ha83/.

The basic steps in ES development are shown in Fig. A1.2. An important stage in ES development is phase 6, rapid prototyping, which includes: knowledge acquisition, knowledge

representation as well as case study identification, testing, and evaluation. Stages of knowledge acquisition are presented in Fig. A1.3.

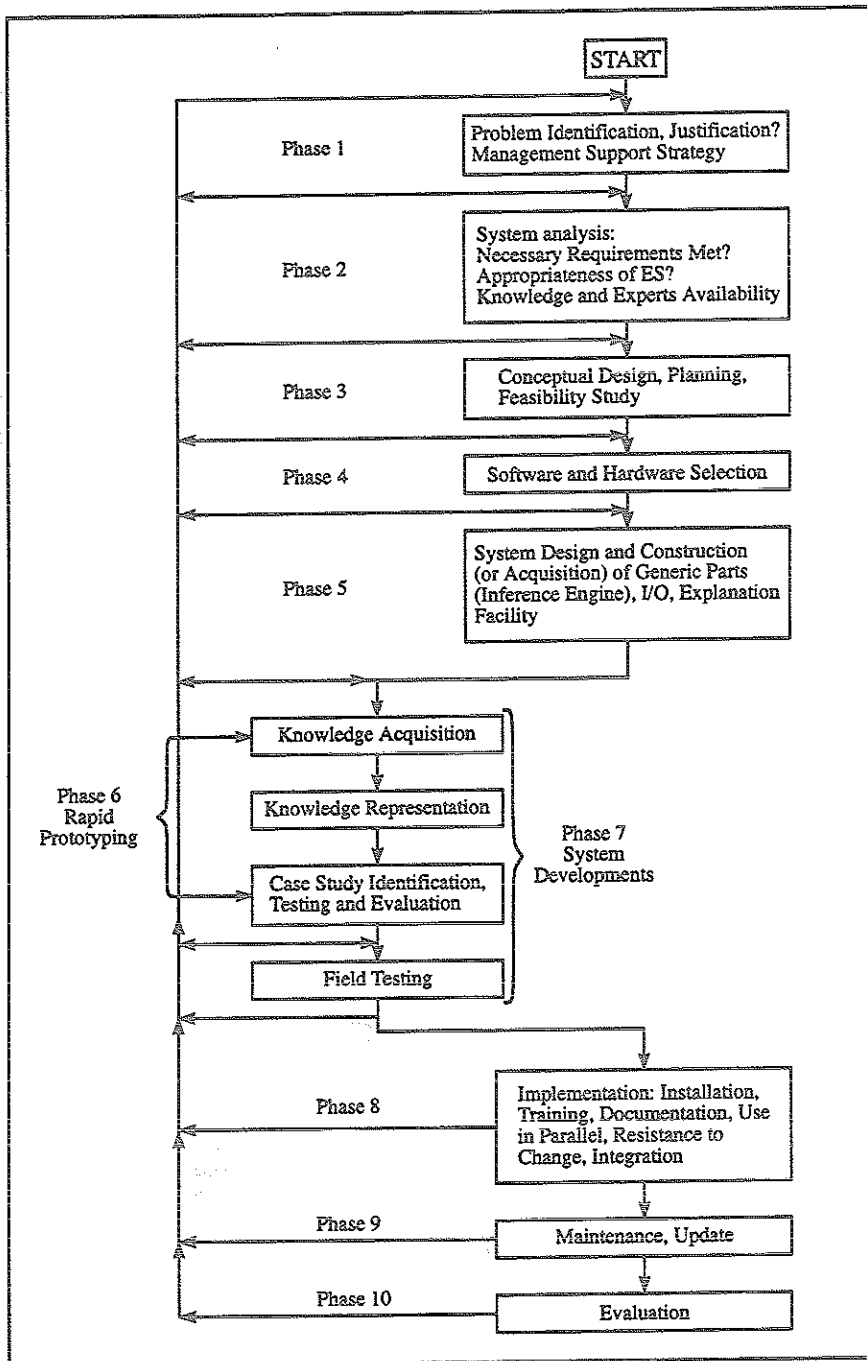


Fig. A1.2: Expert system development process /Tu90/

Knowledge acquisition involves quality control aspects. These appear under the terms "evaluation", "validation", and "verification". The objective of evaluation is to assess an expert system's overall value. In addition to exhibiting acceptable performance levels, ES should be usable, efficient, and cost effective. Validation is a part of evaluation that deals with the performance of the system (e.g., as it compares to the experts). Verification is also part of

evaluation. It refers to building the system "correctly", that is substantiating that the system correctly implements its specification /Tu90/.

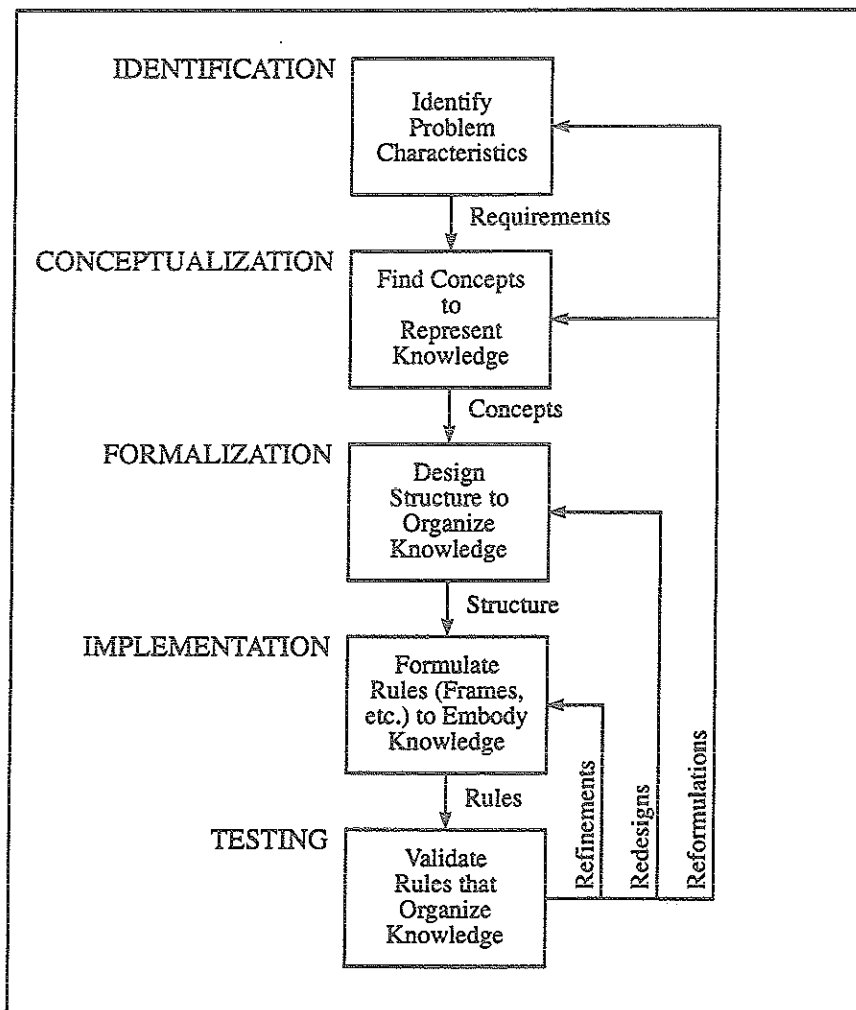


Fig A1.3: Stages of knowledge acquisition /Tu90/

It is helpful to identify several levels of software (and sometimes hardware) that have been involved in ES and their construction. The collection of software can be divided into three major groups: specific expert systems, shells, and tools (tools appear under several names: hybrid environment, knowledge engineering environment, building tools, development tools, etc.) /Tu90/. Building tools can be used to build shells or to build specific ES. ES shells are used to build specific ES (sometimes in conjunction with building tools). These three levels of technology are shown in Fig. A1.4.

The tendency is observed to integrate Decision Support Systems (DSSs) and Expert Systems. DSSs are often understood as the management decision systems, which are defined as interactive computer-based systems to be designed for helping the decision makers to utilize data and models to solve unstructured problems. DSSs couple the intellectual resources of individuals with the capabilities of the computer to improve the quality of decisions. In certain problem domains both ES and DSS may have distinct advantages that, when combined, can

yield synergetic results. Expert systems can be attached to DSS components as illustrated in Fig. A1.5.

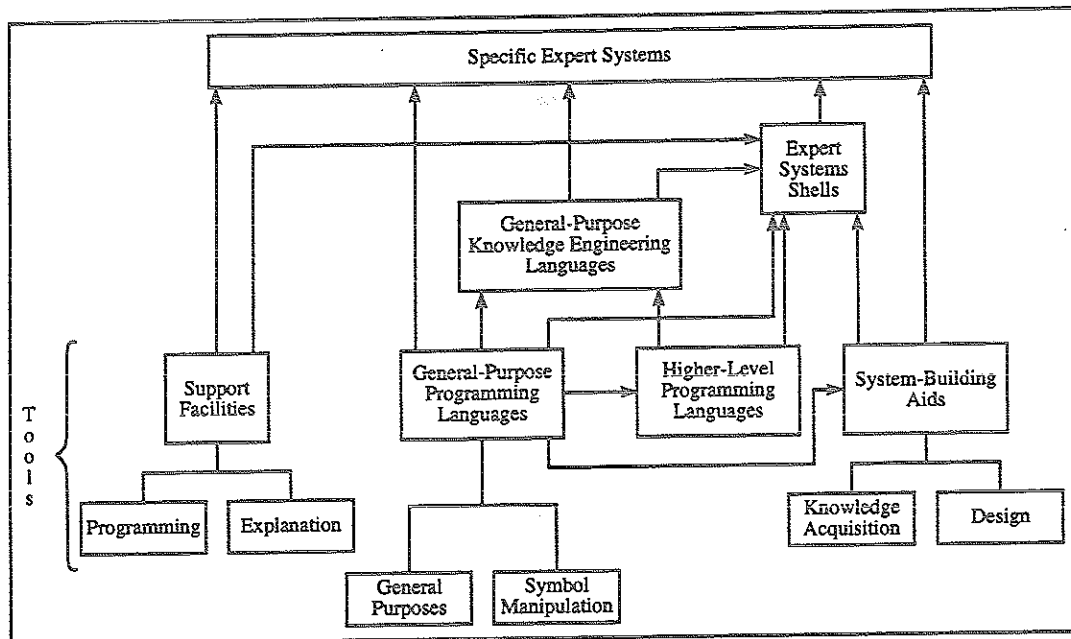


Fig. A1.4: Software for building expert systems /Tu90/

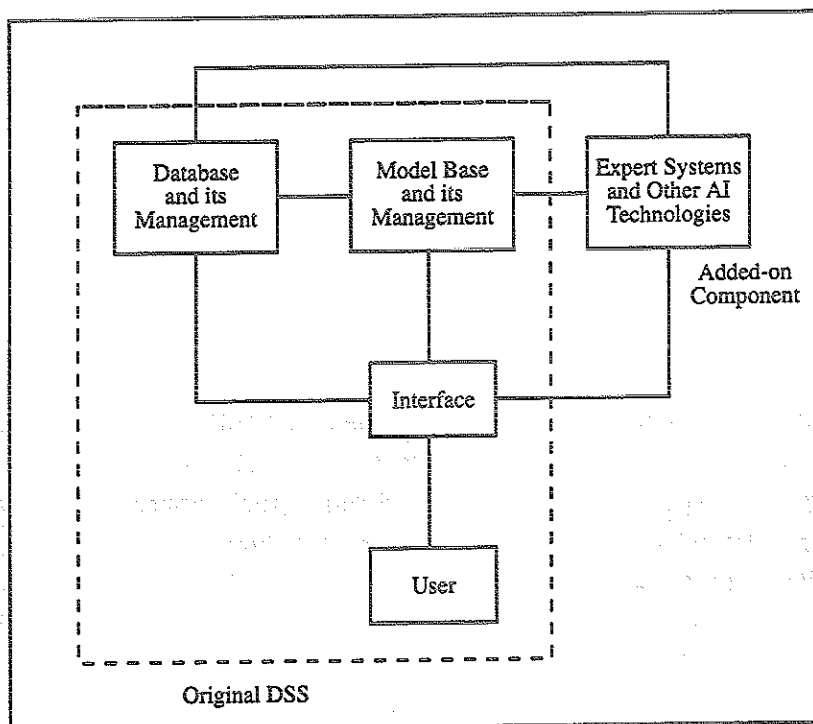


Fig. A1.5: Expert System as a separate component of Decision Support System /Tu90/

A1.2 Software for building engineering expert systems

In recent years a significant number of software tools have arrived on the market for the development of expert systems. As has been shown in Appendix 3 there is a growing interest in developing engineering expert systems on PCs and work stations (e.g. SUN computers), therefore, this short review is devoted to the expert system development tools for such computers, and software tools available in Europe.

Choosing an appropriate tool for building an expert system is a difficult decision. The following considerations should be taken into account when selecting software for the project /Re90b/:

- development constraints (the time, money, personnel and hardware required; there are three main types of tool available: expert system shells, more sophisticated 'hybrid' environments, and high level programming languages such as LISP);
- tool support facilities (useful support facilities include knowledge base editors, debugging aids and explanation facilities; the provision of graphical toolkits can also improve the efficiency of the development process);
- reliability (it is best to select a tool which has a large user community and a reputation for being robust and well-debugged);
- product support/maintenance (select a tool which has adequate documentation, will be maintained/supported by its developer and will support enhancement in other software in use in the future);
- maintainability/updating (as the systems become more complex and larger, significant problems are encountered in maintaining a rule base, validation of the knowledge base and run-time speed; on the other hand, frame-based systems, or object-oriented systems, need to be updated and maintained by a developer who has to be reasonably well trained);
- interfaces (it may be necessary for the expert system to communicate with other hardware and software);
- special requirements (the application may require knowledge to be stored in a certain way, discounting some of the shells available);
- cost (the cost of the tool and associated run-time licences needs to be considered; for most projects these factors are insignificant in comparison to the overall project cost);
- hardware (the expert system to be used needs to be compatible with existing hardware, or the modifications/purchase of hardware to make them compatible needs to be cost-effective);
- standardization (to ensure consistency within a large organization, and the portability and interfacing of applications developed).

Many tools for developing expert systems ranging from small to complex hybrid tools have been described in /Re90b/. Taking into account the needs to develop engineering expert systems for PCs and work stations, to support the human reliability evaluation and PSA studies, the following tools have been selected for further consideration: KEE and KAPPA from IntelliCorp, Inc. as well as EGERIA and XiPLUS from Expertech Ltd. The software products from Intellicorp, Inc. are known worldwide and used for both research and engineering applications.

KEE (Knowledge Engineering Environment) /KE90/, written in CommonLISP, is a hybrid system built on an object-oriented environment to provide knowledge representation tools, various reasoning strategies, and graphical user and developer interface facilities. KEE PC-HOST enables the development of applications on high-powered work stations and the delivery of them onto PCs linked to mainframe/mini-host computers /Re90b/. This tool is quite

popular in developing engineering complex expert systems and therefore is described in detail in Appendix 2.

The approximate prices of KEE are: 10,000 USD (for IBM-PC), 30,000 USD (for mini-computers) and 98,000 USD (for IBM mainframe).

KAPPA /KA92/, written in C, is a general purpose object oriented software environment for developing PC based applications under MS Windows. KAPPA-PC combines four technologies: object-oriented programming, rule-based reasoning, active graphics and dynamic links to other applications and data bases. The processes of the domain are represented in three different ways: methods, functions and rules. KAPPA takes advantage of computer systems with expanded and extended memory and makes full use of Dynamic Data Exchange (DDE).

KAPPA-PC provides developers: powerful modelling of real-world solutions; fast and efficient object management schemes; true message passing; consistent inheritance by reference of slots, values, methods, and slot options; object networks, implicit value typing, slot value constraints (numerical, Boolean, multiple choice of text or objects as slot values); slot monitors (also referred to as demons; methods attached to slots are triggered upon occurrence of a certain event; if needed, before change, after change); message passing, activation of the inference engine, function calls, graphics animation, and data access from methods.

Rules enable developers to incorporate judgement and experience. Rules and object slots are compiled into an RETE inference network. Rules in KAPPA allow: agenda-controlled forward chaining; goal-driven backward chaining; forward engine can be invoked during backward inferencing and vice versa; four rule firing scheme: depth first, breadth first, best first, and selective); pattern matching on objects; priorities for conflict resolution; rule sets focus inferencing; "if" and "then" parts of the rule to be an arbitrary set of conditions and actions; backward chaining on multiple conclusions; flexible explanation facility to explain the conclusions; rule trace and break capabilities; slot trace and break capabilities; message passing, activation of the inference engine, function calls, graphics animation and data access from rules.

KAPPA-PC Application Language (KAL) is an integrated, high-level language which allows rapid prototyping, produces an efficient code and provides a library of about 300 predefined functions. By using KAL developers can: interpret code for rapidly defining, building, testing, and refining prototypes without stopping to compile and link; generate compiled C code for fast runtime performance; manipulate application knowledge, maths, strings, lists, logic, files, control blocks, windows, pop-up menus, input forms, application graphics, interfaces, and system access; pass messages, and activate the inference engine, function calls, graphics animation, and data access from functions.

KAPPA-PC can interface to data bases, spreadsheets, graphics and CAD packages, conventional programming languages, and ASCII files. By taking advantage of DDE, KAPPA-PC provides an interface to other DDE applications, such as Asymetix Toolbook, MS Excel as well as Word and Project for Windows. KAPPA can dynamically interface to other programmes written in C, FORTRAN, or PASCAL. Embedding a C, FORTRAN, or PASCAL function in KAPPA is accomplished by declaring it in KAPPA's main module, compiling, and linking. Any executable code can be called from KAPPA. KAPPA can be embedded into any other MS Windows application. The approximate price of KAPPA-PC is 4,000 USD.

Recently ProKAPPA has become available, a general purpose application development environment for UNIX, and X Windows work stations. ProKAPPA is also written in C and is considered as a very powerful hybrid shell, comparable functionally with KEE, but faster. Intellicorp, Inc. announced recently that a new generation of software for building expert systems will be available in the near future. Ideas from two shells, namely KAPPA 2.0 and ProKAPPA, will be further developed based upon a new object-oriented programming methodology. This new hybrid tool, named KAPPA 3.0, for unix work stations will be a successor to KEE.

EGERIA /Re90b/, written in C, is a hybrid tool which supports classes, inheritance, objects and relationships for knowledge representation. The software and hardware environments for applying this tool range from IBM-PC or IBM PS/2 to SUN or VAX workstations. EGERIA provides the following facilities for knowledge representation: object description (frames, multiple inheritance, objects, relationships, rules), data types (conditions, probabilities, numbers, enumeration, string, groups, functions) and actions (tasks, procedures, messages).

The basic inferencing mechanism of EGERIA is forward chaining, with the propagation of consequences through the inference net established by the compiler. Backward chaining can be forced by user-specification of goals during run-time. Backward chaining carried out by EGERIA is 'intelligent' in the sense that it minimizes unnecessary questions to the user.

In EGERIA the inference engine handles: forward chaining, intelligent backward chaining, hypothetical reasoning (viewports, truth maintenance, hypothesize and test), object oriented, blackboard, logic (first order, set theory), production rules, demons, metacontrol, uncertainty management (Bayesian, probabilistic, fuzzy sets, multivalues), pattern matching (variables, literals, strings, structures), maths calculation and modularization.

EGERIA provides a built-in data base, which is dBASE III compatible on the PC. Because it is controlled from the procedural language, the data base I/O is well integrated with the knowledge base. A direct interface to C is available; external programmes can be run through the C interface.

The approximate prices of EGERIA are: 10,000 USD (for IBM-PC or compatible) and 15,000-44,000 USD (for SUN and VAX work station).

XiPLUS /Re90b/, written in C (since release 3), is a PC shell in which rules are represented in English language format. Rules are made up of identifiers, relations and values. This shell is fundamentally backward-chaining, although limited forward chaining is supported by the use of demons. XiPLUS provides no support for uncertainty handling. The shell allows for programme interfaces to C, Assembler and Basic. There are also data interfaces to spreadsheets and data base files and graphics interfaces to graphics packages. An interface programme ROLLPROG allows XiPLUS to be dumped to a disk file making more memory available for an external application.

The approximate prices of XiPLUS are: 4,000 USD (for IBM-PC) and 8,000-34,000 USD (for VAX work station).

Appendix 2

A hybrid software tool for building expert systems:

Knowledge Engineering Environment (KEE)

A2.1 General remarks

KEE is a software package which provides the user with the programming tools and techniques necessary to generate an application representing and analysing knowledge (expert system). A model of the problem to be handled is represented by a set of hierarchically ordered objects and a set of rules. Properties can be defined for these objects and their behaviour can be simulated by programs or conclusions on the basis of these rules.

KEE is written in Lisp and more extensive applications such as the system of rules are themselves implemented as a KEE expert system.

A2.2 The structure of KEE

A2.2.1. Units

Objects are represented in KEE by units. KEE differentiates between a class unit (a set of similar objects) and a member unit (a representative of such a class unit). The classification of objects into class units or member units must be made by the person generating the expert system and depends on the desired degree of detail definition.

Example 1:

The following classification could be made for a problem from reactor technology:

class unit: active.components
member unit: valve

or

class units: active.components
 valves
member units valve.with.electric.drive
 valve.with.hydraulic.drive

The second variant is characterized by a higher degree of detail definition.

The units are hierarchically ordered. The linkage between two class units is referred to as superclass-subclass link (shown by a solid line), the linkage between class unit and member unit is a class-member link (shown by a dotted line). There are no limits to the combination of these links, so that a unit can have several superclasses or be linked to another unit by a superclass-subclass link and to a further one by a class-member link.

Not all units of an expert system must be interlinked. Such a system normally consists of several hierarchical unit structures. The classification and hierarchical order of units can be changed and extended at any time.

One also refers to the descendants of a unit within a unit structure. Descendants are all hierarchically lower units.

Example 2:

Figure A2.1 shows such a unit structure, again with terms from reactor technology as the objects.

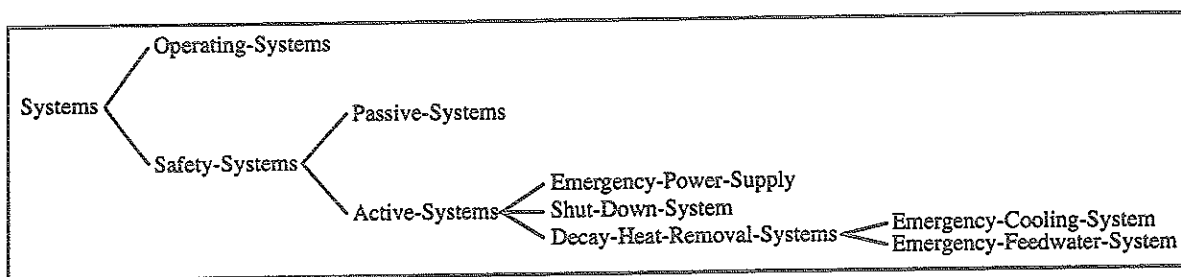


Fig. A2.1: Example of a unit hierarchy with terms from reactor technology

The systems are divided into Operating-Systems and Safety-Systems through superclass-subclass links, and the Safety-Systems are again divided into Active- and Passive-Systems. Two representatives of the Active-Systems (identified by a class-member link) are Emergency-Power-Supply and Shut-Down-System. A group of similar objects belonging to the Active-Systems (superclass-subclass link) are the Decay-Heat-Removal-Systems with Emergency-Cooling-System and Emergency-Feedwater-System as representatives (class-member link).

A2.2.2 Slots

Slots are defined for each unit to describe an object. A slot can be a property (attribute) of the object (e.g. colour) and also its behaviour (e.g. a calculation rule).

KEE differentiates between two groups of slots, the member slots and the own slots.

An own slot defines a particular property of an object. Member units as the representatives of a class unit have therefore only own slots.

A member slot is the common property of a set of objects, i.e. of a class unit and its descendants. All descendants inherit the member slot from the class unit in which it is defined. If the descendant is a member unit, the member slot automatically becomes an own slot. A class unit can also possess own slots in addition to member slots, but these are not inherited.

It is advisable to define member slots as high as possible in the hierarchy.

Example 3:

The hierarchy in Fig. A2.1 is extended by several slots.

Unit	Slots defined in this unit	
	member	own
Systems	maintenance.interval reliability	
Safety-Systems	degree.of.redundancy	
Active-Systems	power.supply external.actuation	action.mode

Tab. A2.1: Slots defined in the hierarchy of Fig. A2.1

In addition to the slots defined in the unit active.systems, the latter would inherit the slots maintenance.interval and reliability from the unit Systems and the slot degree.of.redundancy from the unit Safety-Systems. The member unit Shut-Down-System (Fig. A2.1) has the slots maintenance.interval, reliability (inherited from Systems), degree.of.redundancy (inherited from Safety-Systems), power.supply and external.actuation (inherited from Active-Systems). The slot action.mode is an own slot and therefore not inherited. Since Shut-Down-System is a member unit, all its slots are own slots.

A2.2.3 Facets

Each slot consists of a set of facets describing it in more detail. The most important facets are inheritance, valueclass and value.

The value facet contains the value of the slot.

The inheritance facet determines how the contents of the value facet are transmitted for the inheritance of slots. There are several possibilities, such as override.values (explicitly entered values overwrite the inherited value) or union (both the entered value and the inherited value are values of the value facet).

The valueclass facet defines the valid range of values for the value facet. Examples of valueclasses are number domains such as INTEGER or NUMBER, an interval, a group of values such as (blue, red, green), a class unit with all its member units as the valid values or a combination of these possibilities.

The definition of a class unit as a valueclass is a further type of linkage between units in addition to the links described above.

Example 4:

Terms and notations from reactor technology are used again.

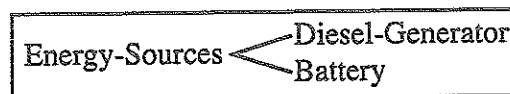


Fig. A2.2: Unit hierarchy of energy sources

In addition to the hierarchy in Figure A2.1 there is now a second hierarchy which models possible power sources. The following slots and facets are defined within the units of Fig. A2.1:

class unit: Active-Systems
 member slot: power.supply
 valueclass: Energy-Sources

member unit: Emergency-Feedwater-System
 own slot: power.supply
 value: Diesel-Generator

The unit Active-Systems from Fig. A2.1 has the slot power.supply as a common property which is characterized by the source. Since the power sources are organized in a hierarchy of their own (Fig. A2.2), the valueclass facet receives the value Energy-Sources. In the member unit Emergency-Feedwater-System the inherited slot power.supply becomes an own slot and receives the value Diesel-Generator, a member unit of Energy-Source.

A2.2.4 Methods

The values so far described for the facets can only define the properties of a unit. If the behaviour of an object is to be simulated, a so-called method slot must be defined. A method slot is a normal member or own slot whose inheritance and valueclass facet have the value

'method'. The value facet of this slot contains a Lisp programme. The inheritance of method slots takes place according to the rules described in the preceding chapter. A method slot is started by a send message command.

Example 5:

The example of method slots originates from mathematics. All slots are method slots. They indicate rules according to which the volume of bodies is calculated.

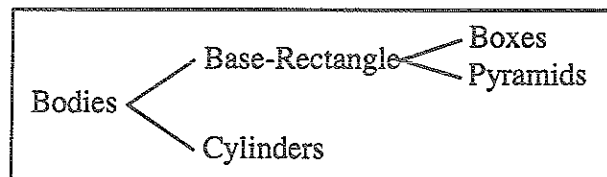


Fig. A2.3: Unit hierarchy of bodies

The following method slots are defined in the units:

unit: Bodies

member slot: calculate.volume

value: volume = (base * height)

unit: Base-Rectangle

member slot: calculate.base

value: base = (width * length)

unit: Pyramid

own slot: calculate.pyramid.height

value: pyramid.height = (1/3 * height)

unit: Cylinder

own slot: calculate.base

value: base = (pi * radius²)

The volume calculation is given for all bodies by the formula base*height and can therefore be defined by a member slot at the top of the hierarchy. The different bases must be distinguished in calculating the base. Bodies was therefore divided into two subclasses, one defining the calculation for a circle and the other that for a rectangle as the base.

If e.g. the volume of a pyramid is calculated, the general formula base*height is obtained from the unit Bodies, the formula for the base from the unit Base-Rectangle and the formula for the height from the unit Pyramid.

A2.2.5 Active values

A further possibility of activating Lisp programmes in slots is the definition of active values. The active value is a facet whose value is a Lisp programme. If the value facet of the slot to which the active value belongs is accessed, the Lisp programme is executed. KEE differentiates between active values for read and write access. In the latter case, the value of the value facet is only changed after execution of the Lisp programme.

Examples of application are the verification of data read in, protection against unauthorized reading or the automatic correction of data in the case of a value change.

Example 6:

The unit systems from Example 3 has the slots maintenance.interval and reliability. The value of the value facet of reliability is i.a. also dependent on the value of the maintenance.interval. In order to achieve data consistency, the following definitions are made for the slot maintenance.interval:

value class: ONE.OF (monthly annually)
active value for write access.

The Lisp programme of the active value executes the following algorithm:

```
if
    the old value of maintenance.interval is monthly and
    the new value of maintenance.interval is annually
then
    the value of reliability is (value * factor)

if
    the old value of maintenance.interval is annually and
    the new value of maintenance.interval is monthly
then
    the value of reliability is (value / factor)
```

This active value thus serves to automatically adapt the value of reliability to a change in maintenance.interval.

A2.2.6. Rules

A second approach to simulating the behaviour of objects is the use of rules. A detailed description of the rules would be going too far within this context, so that in the following some important properties of the rules will be enumerated and their behaviour demonstrated using an example.

The general form of a rule is:

IF premises
THEN conclusion.

A rule is the value of the value facet of a rule unit. The definition of a rule unit automatically creates a superclass link to an expert system available in KEE which organizes the processing and linking of rules.

Rules are formulated in a language similar to English in order to make them comprehensible even for inexperienced users.

Rules can be combined in classes (rule class units), so that the class can be specified for the processing of rules.

KEE offers two fundamental possibilities of logical conclusions by means of rules: forward chaining and backward chaining. Forward chaining begins with a new premise and determines the resultant conclusions, whereas backward chaining uses known facts in order to verify a hypothesis.

A2.2.6.1 Forward chaining

The forward chaining mode will be explained by an example.

Example 7:

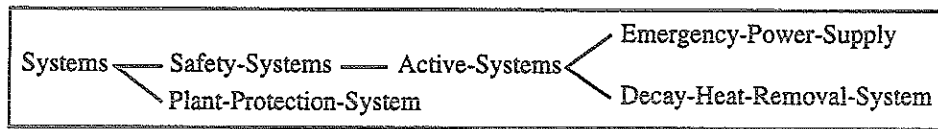


Fig. A2.4: NPP system in a unit hierarchy

The unit structure of Fig. A2.1 has been extended by the unit Plant-Protection-System. Fig. A2.4 only shows the units of relevance for this example.

The following slots and facets are defined:

unit: Plant-Protection-System

own slot: loss.of.offsite.power

valueclass: ONE.OF (yes no)

own slot: emergency.protection

valueclass: ONE.OF (yes no)

unit: Active-Systems

member slot: demand

valueclass: ONE.OF (yes no)

member slot: availability

valueclass: ONE.OF (yes no)

The next figure shows the ruleclass simulation.rules with three rules.

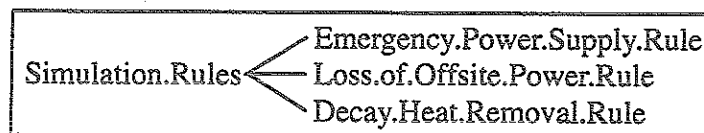


Fig. A2.5: Unit hierarchy of simulation rules

The rules are defined as follows:

Loss.of Offsite.Power.Rule:

IF value of loss.of.offsite.power in Plant-Protection-System is 'yes'

THEN value of demand in Emergency-Power-Supply is 'yes'

Emergency.Power.Supply.Rule

IF value of demand in Emergency-Power-Supply is 'yes' and

value of availability in Emergency-Power-Supply is 'no'

THEN value of availability in Decay-Heat-Removal-System is 'no'

Decay.Heat.Removal.Rule:

IF value of availability in Decay-Heat-Removal-System is 'no'
 THEN value of emergency.protection in Plant-Protection-System is 'yes'

The following initial values are set in the expert system:

Unit	Slot	Value
Plant-Protection-System	loss.of.offsite.power	no value
	emergency.protection	no value
Emergency-Power-Supply	demand	no value
	availability	no
Decay-Heat-Removal-System	availability	no value

Tab. A2.2: Slot values before rule application in forward chaining example

Processing of the rule system by forward chaining comprises the following steps:

- Start of forward chaining by the command ASSERT statement ruleclass
 statement: *value of loss.of.offsite.power in Plant-Protection-System is 'yes'*
 ruleclass: simulation.rules
- The statement is added to the expert system as new knowledge.
- Seek all rules in the ruleclass which have the statement as the premise
- The rule Loss.of.Offsite.Power.Rule is found
- The conclusion *value of demand in Emergency-Power-Supply is 'yes'* is added to the expert system as new knowledge
- Forward chaining is restarted with the conclusion as the new statement
- The rule Emergency.Power.Supply.Rule is found
- The second premise *value of availability in emergency.power.supply is 'no'* is verified
- The premise is fulfilled (see Tab. A2.2)
- The conclusion *value of availability in Decay-Heat-Removal-System is 'no'* is added to the expert system as new knowledge
- Forward chaining is restarted with the conclusion as the new statement
- The rule Decay.Heat.Removal.Rule is found
- The conclusion *value of emergency.protection in Plant-Protection-System is 'yes'* is added to the expert system as new knowledge
- Forward chaining is restarted with the conclusion as the new statement
- No rule is found
- Forward chaining is terminated

After the execution of these steps, the values in the expert system have changed as follows:

Unit	Slot	Value
Plant-Protection-System	loss.of.offsite.power	yes
	emergency.protection	yes
Emergency-Power-Supply	demand	yes
	availability	no
Decay-Heat-Removal-System	availability	no

Tab. A2.3: Slot values after rule application in forward and backward chaining example

A2.2.6.2 Backward chaining

Backward chaining is also merely demonstrated by an example.

Example 8:

The unit hierarchy, the slots and facets, and the rule system are defined as in example 7. The initial values are shown in Tab. A2.4.

Unit	Slot	Value
Plant-Protection-System	loss.of.offsite.power	yes
	emergency.protection	no value
Emergency-Power-Supply	demand	ye
	availability	no
Decay-Heat-Removal-System	availability	no value

Tab. A2.4: Slot values before rule application in backward chaining example

Processing of the rule system by backward chaining comprises the following steps:

- Start of backward chaining by the command QUERY hypothesis ruleclass
hypothesis: *value of emergency.protection in Plant-Protection-System is 'yes'*
ruleclass: Simulation.rules
- Seek all rules in the ruleclass which have the hypothesis as the conclusion
- The rule Decay.Heat.Removal.Rule is found
- The premise *value of availability in Decay-Heat-Removal-System is 'no'* is added to the expert system as new knowledge
- Backward chaining is restarted with the premise as the new hypothesis
- The rule Emergency.Power.Supply.Rule is found
- The first premise *value of demand in Emergency-Power-Supply is 'yes'* is verified
- The second premise *value of availability in Emergency-Power-Supply is 'no'* is verified
- The premise is fulfilled (see Tab. A2.3)
- There is no further hypothesis to be examined

- The first hypothesis *value of emergency protection in Plant-Protection-System is 'yes'* is added to the expert system as new knowledge
- Backward chaining is terminated

After the execution of these steps, the values in the expert system are as described in Table A2.3.

Appendix 3

An expert system for the application of THERP

A3.1 Introduction

The expert system ESAP presented in this Appendix is intended to increase the application reliability of the THERP method (see Chapter 3.5.2). The suitability of an expert system for this purpose is based on

- easy expandability
- a knowledge base growing with application
- the availability of the required data management and
- ease of creating a user-friendly display surface.

Chapter 2 first of all illustrates the content conversion of THERP. Chapter 3 describes the program conversion using KEE (see Appendix 2).

A3.2 Content conversion of THERP

A3.2.1 Procedure

THERP provides a model for calculating the probability with which human errors contribute to a system failure by analysing the procedures to be carried out by the maintenance staff and operators. THERP distinguishes between two times at which a procedure takes place: prior to or during an accident. Procedures prior to an accident serve for component testing, maintenance and timely repair, if required, and thus contribute towards ensuring system reliability. Procedures during an accident concern the supervision of automatically initiated system responses, planned measures for accident control and the implementation of emergency measures for beyond-design events.

The example under consideration is the safety concept of the FRJ-2 (DIDO) research reactor. It provides for two possibilities of emergency cooling for the accident involving a "leak in the primary circuit": (1) automatically initiated heavy water emergency cooling; (2) operator-initiated light water emergency cooling (Fig. A3.1).

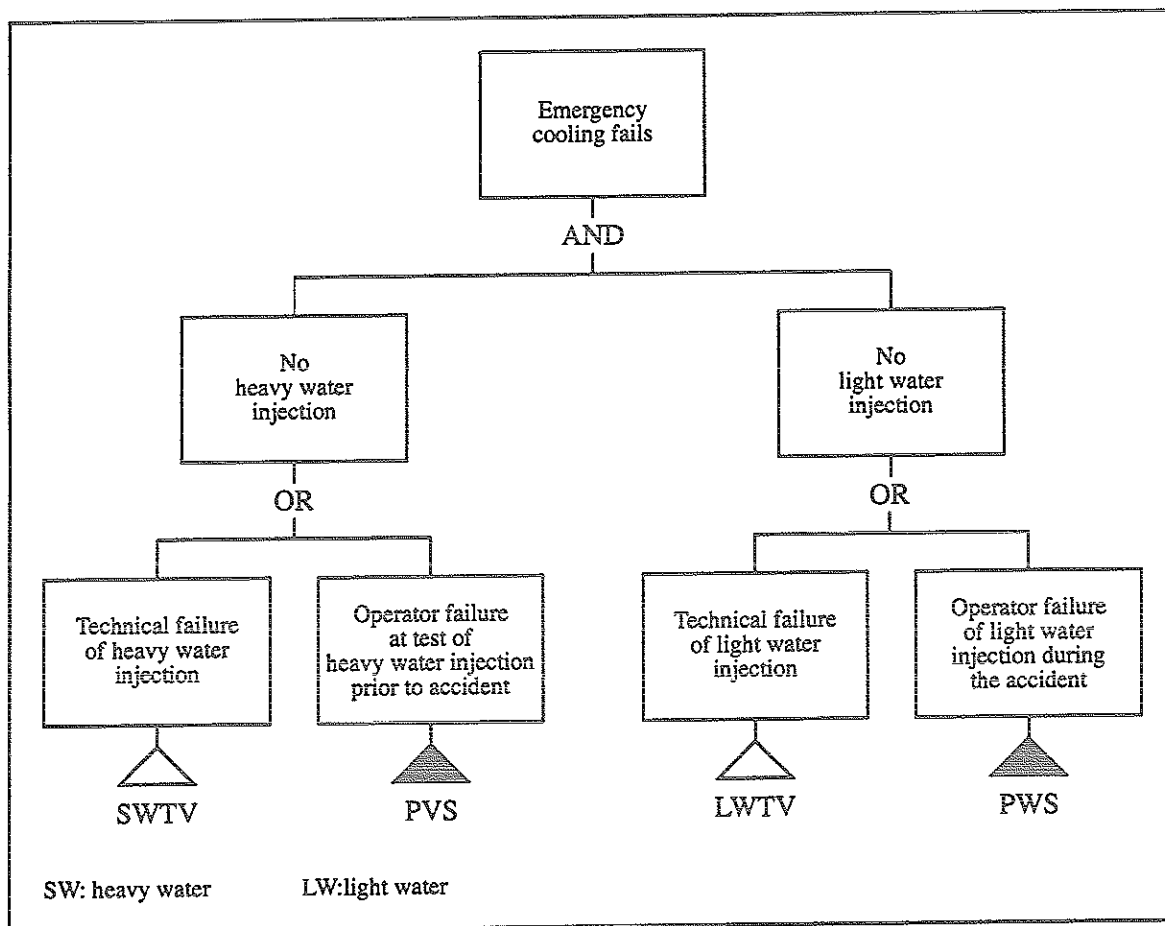


Fig. A3.1: Simplified fault tree for FRJ-2 emergency cooling

In order to evaluate the operator error contribution to the system failure "emergency cooling fails" it is necessary to calculate the probabilities of the occurrence of a critical fault in either

procedure: "heavy water injection test (every 4 years)" (procedure prior to the accident) or "light water injection (on demand)" (procedure during the accident).

The calculation of these probabilities with THERP requires an extensive analytical process (note: the THERP handbook contains about 800 pages). In order to assist the user in this analysis, a catalogue of questions to be answered for evaluation was devised as part of the studies described here. For example, the operator's knowledge level or a possible staff redundancy must be confirmed by a 2nd operator. The questions for procedures prior to an accident are compiled in Tab. A3.1. Questions for procedures during an accident are comparable. The questions are written in standard type, the possible answers in italics.

The procedure is performed by an...				
<i>experienced operator</i>		<i>inexperienced operator</i>		
The procedure is known to the operator due to...				
<i>an oral instruction</i>	<i>a written document</i>			
	Is the written document available as a check list?			
	<i>yes</i>		<i>no</i>	
	How is the procedure described			
	<i>Activity within the control room</i>	<i>Work at valves</i>	<i>Test, calibration</i>	<i>Maintenance</i>
At which level of stress is the procedure performed?				
<i>too low</i>	<i>optimal</i>	<i>high</i>	<i>extremely high</i>	
Personnel redundancy?				
<i>no</i>	<i>by 2nd operator</i>		<i>by 2nd and 3rd operator</i>	
	Control situation for the 2nd operator			
	<i>Routine control</i>	<i>Special control</i>	<i>Control with active manual measure</i>	<i>Undetected error = peril of death</i>

Tab. A3.1: Questions concerning the procedure prior to an accident

The individual steps are then to be listed for each of these procedures. All steps or combinations of steps which, if omitted, lead directly or indirectly to a system failure must be identified. The individual steps and step combinations thus found constitute the set of so-called "failure paths". Several such failure paths are considered as a unit, if they are completely interdependent in the sense of forming a definable independent complex of tasks. One failure path is selected from these units as representative, and only this path is then used for further calculations.

The cognitive level characterizing the type of activity must then be determined for the steps resulting from the above procedure. A distinction is made between high (knowledge-based), medium (rule-based) and low (skill-based) level.

For each combination of steps to be included in the calculation it is furthermore necessary to define the dependence between two successive steps. This dependence is divided into none, low, medium, high and total.

The first four steps in the "light water injection" procedure are, for example, "take procedure in hand", "give orders for light water injection", "go to emergency control room" and "open cabinet". The omission of one of these steps leads directly to system failure so that all steps thus belong to the set of failure paths. However, these steps are considered as a unit ("complex of tasks") of which step 4 "open cabinet" is representative.

Errors of omission and commission, (perception, operation, and other errors) are considered as the causes of failure which may occur during the performance of these steps. The erroneous performance of a step is always based on one or several of these causes.

Errors of omission (in the sense of forgetting something) are accounted for by previous information concerning the procedure. Additional information is required to include errors of commission. The question pattern for perception errors (e.g. reading) is shown in Tab. A3.2 - Tab. A3.4. Operating errors (e.g. selection) are treated similarly. For the error type "other errors" a more detailed description must be given by the model user himself.

Requires the performance of the step to recognize the information signaled by a reading board?				
<i>no</i>	<i>yes</i>			
	Is it possible not to perform the step by watching the wrong reading board?			
	<i>no</i>	<i>yes</i>		
		Is the considered reading board positioned within a group of similar boards?		
		<i>no</i>	<i>yes</i>	
			How is the considered reading board distinguished from adjacent reading boards	
		<i>only by labelling</i>	<i>additionally by arrangement in a functional group</i>	<i>additionally by clearly drawn separating lines</i>
	Is the performance of a step required to recognize qualitative or quantitative information?			
<i>quantitative</i>		<i>qualitative</i>		
Go on with Tab. A3.3		Go on with Tab. A3.4		

Tab. A3.2: Questions for perception errors

Which description suits the required quantitative recognition best?
<i>Correct reading of a value on ...</i> - a printer with many parameters - a graphical presentation - a sheet writer - an analogue display - a digital display (< 5 digits)

Tab. A3.3: Continuation of Tab. A3.2

Which description suits the required qualitative recognition best?
<i>Noticing a deviation from the limit on ...</i> - a sheet writer without limit lines - a sheet writer with limit lines - an analogue display without limit lines - an analogue display with limit lines - a digital display <i>Correct interpretation of the signal of an indication lamp</i> <i>Noticing a wrong reading</i>

Tab. A3.4: Continuation of Tab. A3.2

For modelling in THERP the following steps must thus be performed:

- identification of procedures belonging to the system failure investigated
- general information about each procedure
- decomposition of the procedure into steps
- identification of essential steps
- information about possible types of errors.

A3.2.2 Calculations

A3.2.2.1 General

The calculation of failure probabilities leading to a system failure will be described in the following. Failure probabilities for individual procedures are calculated, which are dependent on the probabilities of the failure paths or steps belonging to these failure paths.

The step failure probabilities are calculated using the so-called basic failure probabilities (BF) of the SWAIN handbook. A basic failure probability is not assumed to be a fixed parameter, but a random lognormal-distributed quantity. A lognormal distribution is uniquely determined by two parameters. The two parameters specified in the SWAIN handbook for the basic failure

probabilities are median (M) and uncertainty factor (U) for the two-sided 50 % uncertainty interval (M/U, M·U).

The other parameters can be calculated using the formulas given under (1) in Fig. A3.2. In order to calculate the probability of a failure path, a procedure or system failure, it is necessary to use the operation rules specified under (2).

The notations used in this section are always composed of the parameter abbreviation (see Fig. A3.2) and the abbreviations for the probability just calculated. If the value depends on a failure cause, this cause is indexed. The notation MBF_U , for example, denotes the median of the basic failure probability for the failure cause "Unterlassung" (omission).

1. Conversion

Lower bound $UG = \frac{M}{U}$

Upper bound $OG = M \cdot U \quad (\text{max. } 1)$

Variance $V = E^2 \cdot \left(\frac{E^2}{M^2} - 1 \right)$

Median $M = UG \cdot U = \frac{E}{\sqrt{1 + \frac{V}{E^2}}}$

Uncertainty factor $U = \frac{OG}{M} = \frac{M}{UG}$

Mean $E = M \cdot \exp \left(0.5 \cdot \frac{\{\ln(OG / UG)\}^2}{3.29^2} \right)$

2. Combining of n events

AND Operation:

Median $M = \prod_{i=1 \dots n} M_i$

Lower bound $UG = \frac{M}{\exp \left(0.5 \cdot \sqrt{\sum \{\ln(OG_i / UG_i)\}^2} \right)}$

Higher bound $OG = M \cdot \exp \left(0.5 \cdot \sqrt{\sum \{\ln(OG_i / UG_i)\}^2} \right)$

OR Operation:

Mean $E = 1 - \prod_{i=1 \dots n} (1 - E_i) \approx \sum E_i$

Variance $V = \sum V_i$

Fig. A3.2: Computation rules for lognormal-distributed probabilities

A3.2.2.2 Basic failure probabilities

First of all, the medians and uncertainty factors of the basic failure probabilities (BF) are determined for each step (cf. A3.2.1) and each type of occurring error, i.e. U_1 , U_2 and U_3 for omission errors, B_1 and B_2 for operating errors, W_1 and W_2 for perception errors and S for other errors.

The following definitions apply:

- U_1 : EOM with correct use of checklist
- U_2 : EOM with incorrect use of checklist
- U_3 : EOM since written documents are not used
- B_1 : ECOM because components were mixed up or wrongly plugged together
- B_2 : ECOM because sense of rotation was confused
- W_1 : ECOM because displays were mixed
- W_2 : ECOM in reading

The basic failure probabilities for omission errors are thus dependent on the general information about the procedure and the number of steps belonging to the procedure. Tab. A3.5 shows an example of determining the median and uncertainty factor of the basic failure probability of U_1 (MBF_{U_1} , UBF_{U_1}).

The procedure is known to the operator due to an oral instruction (see Tab. A3.1)					The procedure is known to the operator due to a written document (see Tab. A3.1)	
$i=1^1$	$i=2$	$i=3$	$i=4$	$i=5$	$n \leq 1^2$	$n \geq 1$
0.001 (3)	0.003 (3)	0.01 (3)	0.03 (5)	0.1 (5)	0.001 (3)	0.003 (3)

¹ step number in the set of important steps (for step numbers >5 a written procedure is assumed)

² total number of steps

Tab. A3.5: Median (uncertainty factor) of the basic failure probability of U_1

The basic failure probabilities for perception errors, operating errors and other errors are dependent on the type of procedure (prior to or after the accident), on the cognitive level belonging to the step and on the answers to the specific questions concerning the causes of failure. As an example, the determination of the median and uncertainty factor of the basic failure probability of W_1 (MBF_{W_1} , UBF_{W_1}) is shown in Tab. A3.6 (page A4-8).

The basic failure probabilities of W_2 , B_1 , B_2 and S are determined by similar rules.

A3.2.2.3 Modified failure probabilities

After determining the basic failure probability, its medians and uncertainty factors must be modified so that modified failure probabilities (MF) are obtained. These modifications are dependent on the general information about the procedure, the cognitive level of the step and the specific information about the failure cause.

The performance of the step does <i>not</i> require the information signaled by a reading board to be recognized (see Tab. A3.2)			0 (1)
It is <i>not</i> possible not to perform the step by watching the wrong reading board (see Tab A3.2)			0 (1)
The considered reading board is <i>not</i> positioned within a group of similar boards (see Tab. A3.2)	Procedure prior to accident		0 (1)
	Procedure during accident	The cognitive level is <i>high</i> (<i>knowledge</i>)	0.001 (10)
		The cognitive level is <i>medium</i> (<i>rule</i>)	0.001 (5)
		The cognitive level is <i>low</i> (<i>skill</i>)	0.001 (5)
The considered reading board is distinguished from adjacent reading boards <i>only by labelling</i> (see Tab. A3.2)			0.003 (3)
The considered reading board is distinguished from adjacent reading boards <i>additionally by arrangement in a functional group</i> (see Tab. A3.2)			0.001 (3)
The considered reading board is distinguished from adjacent reading boards <i>additionally by clearly drawn separating lines</i> (see Tab. A3.2)			0.0005 (10)

Tab. A3.6: Median (uncertainty factor) of the basic failure probability of W_1

For example, the modification of U_1 depends on the cognitive level, the stress level and the operator's experience (see Tab. A3.1). These calculations are shown as an example of the median of U_1 in Tab.A3.7.

Stress level											
<i>too low</i>	<i>optimal</i>			<i>high</i>				<i>extremely high</i>			
	Operator			Operator				Operator			
	<i>experienced</i>	<i>inexperienced</i>		<i>experienced</i>		<i>inexperienced</i>		<i>experienced</i>		<i>inexperienced</i>	
		cognitive level		cognitive level		cognitive level		cognitive level		cognitive level	
		<i>high</i>	<i>me- dium low</i>	<i>high</i>	<i>me- dium low</i>	<i>high</i>	<i>me- dium low</i>	<i>high</i>	<i>me- dium low</i>	<i>high</i>	<i>me- dium low</i>
×2	none	×2	none	×5	×2	×10	×4	=0.25	×5	=0.5	×10

Tab. A3.7: Modification of the median of the basic failure probability of U_1

For an optimum stress level, an experienced operator and a high cognitive level the median of the modified basic failure probability for U_1 would thus be:

$$MMF_{U_1} = MBF_{U_1} \times 2$$

A3.2.2.4 Recovery failure probabilities and failure conditions

In the next step the recovery failure probabilities (KF) will be determined, viz. three KFs for procedures prior to an accident (PV_{1a} , PV_{1b} , PV_2) and two KFs for procedures during an accident (PW_1 , PW_2). These correction failure probabilities are dependent on the general information concerning the procedure (see Tab. A3.1). As an example, Tab. A3.8 shows the determination of $KF_{PV_{1a}}$.

No personnel redundancy	Personnel redundancy			
	Control situation for the 2nd operator			
	<i>Routine control</i>	<i>Special control</i>	<i>Control with active manual measure</i>	<i>Undetected error = peril of death</i>
1 (1)	0.1 (5)	0.05 (5)	0.01 (5)	0.001 (5)

Tab. A3.8: Median (uncertainty factor) of a correction failure probability for procedures prior to an accident

Furthermore, the probability of the failure conditions "incorrect use of checklist" (FB_1) and "nonuse of written documents" (FB_2) is required for further calculations. These values are also dependent on the general information about the procedure.

A3.2.2.5 Error type contributions and step failure probability

The probability of the error type contributions (FA) is calculated from the modified probabilities, the recovery failure probabilities and the probabilities of the failure conditions using the "AND" operation rule (see Fig. A3.2). This gives the following formulas for the medians:

For procedures prior to the accident and errors of omission:

$$MFA_{U_1} = MMF_{U_1} \cdot MKF_{PV_{1a}} \cdot MKF_{PV_2}$$

$$MFA_{U_2} = MMF_{U_2} \cdot MKF_{PV_{1a}} \cdot MKF_{PV_2} \cdot MFB_1$$

$$MFA_{U_3} = MMF_{U_3} \cdot MKF_{PV_{1b}} \cdot MKF_{PV_2} \cdot MFB_2$$

For procedures prior to the accident and other causes of failure:

$$MFA_{Xa} = MMF_X \cdot MKF_{PV_{1a}} \cdot MKF_{PV_2}$$

$$MFA_{Xb} = MMF_X \cdot MKF_{PV_{1b}} \cdot MKF_{PV_2} \cdot MFB_2$$

with X from W_1 , W_2 , B_1 , B_2 , S

For procedures during the accident and errors of omission:

$$MFA_{U_1} = MMF_{U_1} \cdot MKF_{PV_1} \cdot MKF_{PV_2}$$

$$MFA_{U_2} = MMF_{U_2} \cdot MKF_{PV_1} \cdot MKF_{PV_2} \cdot MFB_1$$

$$MFA_{U_3} = MMF_{U_3} \cdot MKF_{PV_1} \cdot MKF_{PV_2} \cdot MFB_2$$

For procedures during the accident and other causes of failure:

$$MFA_X = MMF_X \cdot MKF_{PV_1} \cdot MKF_{PV_2}$$

with X from W₁, W₂, B₁, B₂, S

For the step of a procedure prior to the accident caused by a "perception error", the probabilities FA_{W1a}, FA_{W1b}, FA_{W2a} and FA_{W2} are thus obtained whose other parameters can be determined using the computing rules in Fig. A3.2.

The probability of the combined contribution of an error type (FA) is calculated using the OR operation. The following formulas are obtained, for example, for the expectation value:

Error of omission:

$$EFA_U = 1 - \prod_{i=1}^3 (1 - EFA_{U_i})$$

Error of perception, operation or other errors:

$$EFA_X = 1 - \prod (1 - EFA_{X_i})$$

$i \in \{1a, 1b, 2a, 2b\}$ bei PV

$i \in \{1, 2\}$ bei PW

and X from W, B, S

This is then used to determine the step failure probability (SF) via the OR operation giving, for example, the expectation value:

$$ESF = 1 - (1 - EFA_U)(1 - EFA_W)(1 - EFA_B)(1 - EFA_S)$$

A3.2.2.6 Failure path probabilities

It is now necessary to first determine the failure path probability (FB) for each failure path. FP = SF is valid for failure paths which only consist of one step. Otherwise, FP is calculated as the AND operation of the conditional SF values within a path, e.g. the expectation value:

$$EFP = \prod_{i=1}^n [A_{i-1,i} + (1 - A_{i-1,i}) \cdot ESF_i] = \prod_{i=1}^n ESF_{i|i-1}$$

n: number of steps in the failure path

$A_{i-1,i}$: dependence between the (i-1)-th and the i-th step in the failure path

The conditional SF values are calculated from the unconditional SF values and the dependences of two successive steps within the path. The calculation of the median of $SF_{i|i-1}$ is shown as an example in Tab. A3.9.

In order to calculate the uncertainty factors it is assumed that the median of the next higher and next lower dependence level is equal to the upper bound and lower bound, respectively, of the

dependence level considered. The uncertainty factor for slight dependence can thus be calculated, for example, by:

$$\begin{aligned}
 USF_{i/i-1} &= OGSF_{slight} / MSF_{slight} && \text{with} \\
 OGSF_{slight} &= \max(MSF_{medium}, OGSF_i) && \text{and} \\
 OGSF_i &= USF_i \cdot MSF_i && (\text{s. Fig. A3.2}) \\
 MSF_{slight} &= \frac{1 + 19 MSF_i}{10}, MSF_{medium} = \frac{1 + 6 MSF_i}{7} && (\text{s. Tab. A3.9})
 \end{aligned}$$

i=1	i>1				
	Dependence				
	none	slight	medium	high	total
$=MSF_i$	$=MSF_i$	$\frac{1 + 19 MSF_i}{10}$	$\frac{1 + 6 MSF_i}{7}$	$\frac{1 + MSF_i}{2}$	$=1$

Tab. A3.9: Median of the conditional step failure probability $SF_{i/i-1}$

A3.2.2.7 Procedure failure probabilities and system failure probability

The contribution of the analysed procedure to the system failure (procedure failure probability, PF) is obtained as an OR operation of the failure path probabilities; e.g. the expectation value from

$$EPF = 1 - \prod_{i=1}^m (1 - EFP_i)$$

m: number of failure paths

No generally valid formula can be specified for the system failure probability (SA) since it depends on the structure of the fault tree. The calculations are therefore only carried out for the individual procedures. The system failure option only serves for the allocation of procedures.

A3.2.2.8 Summary

The following calculations must therefore be performed:

For each step:

- determine the basic failure probabilities (BF) of the types of error classified as possible
- calculate the modified failure probability (MF) from BF
- determine the recovery failure probabilities (KF) and failure conditions (FB)
- calculate the error type contributions (FA) for each cause of failure from MF, KF and FB
- calculate the step failure probability (SF) from FA

For each path:

- calculate the failure path probability (FP) from SF

For each procedure:

- calculate the procedure failure probability (PF) from FP

A3.3 Programme conversion of THERP into KEE

A3.3.0 General

It will be shown in the following how the methodological approaches described in Chapter 2 and the calculations are converted into a programme using the expert system shell KEE. The system failures are represented as a KEE unit structure. The characteristics of the individual units considered (system failure, procedure and step) are stored in slots. The probabilities are calculated by the KEE rule system and by LISP programmes. A menu control and the input/output are also realized by LISP programmes. All KEE terms used in this section are described in detail in Appendix 2.

A3.3.1 Modelling the structure

A3.3.1.1 Units

Each system failure, each procedure and each step are represented in KEE as a unit. Since a system failure consists of several procedures which are again divided into a number of steps, the hierarchy shown in Fig. A3.3 is obtained for a system failure.

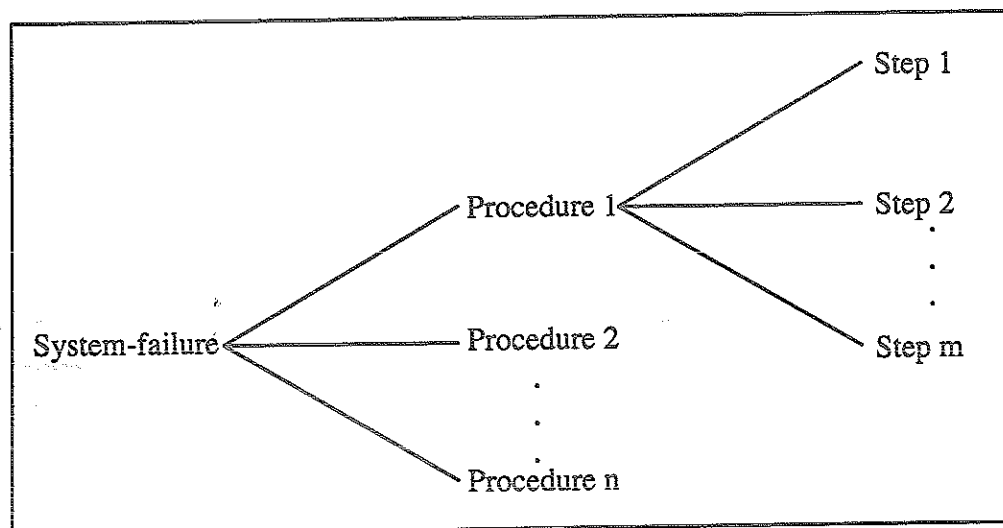


Fig. A3.3: Hierarchy for a system failure

The procedures are divided into procedures prior to the accident (PV) and procedures during the accident (PW). In order to avoid this subdivision into two different types of procedure for each specific system failure, the class unit "Procedures" was introduced with the division into class unit "Procedures-prior-to-accident" and class unit "Procedures-during-accident". Each procedure has a connection to the unit of the associated system failure and to the corresponding procedure class. The subdivision of unit class "Procedures" is shown in Fig. A3.4.

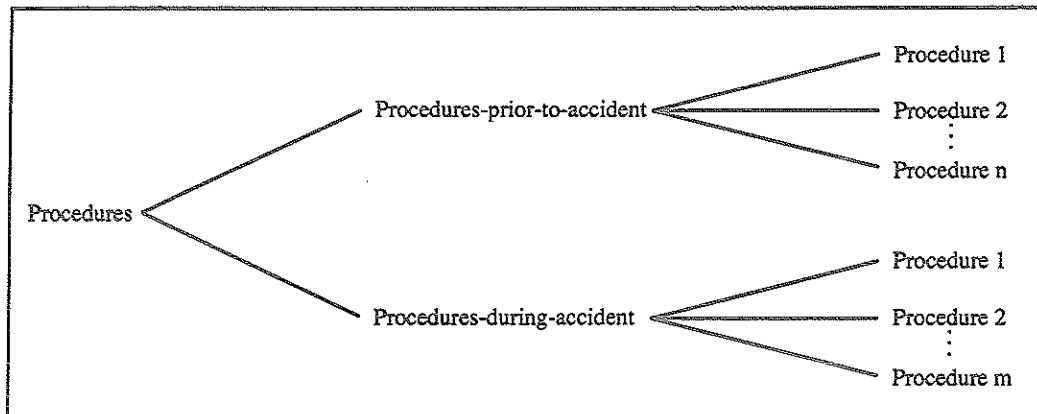


Fig. A3.4: Subdivision of unit class "Procedures"

For reasons of software engineering it is necessary to identify all units for system failures and all units for steps. For this reason, the class units "systems" and "steps" were introduced giving the classification shown in Fig. A3.5.

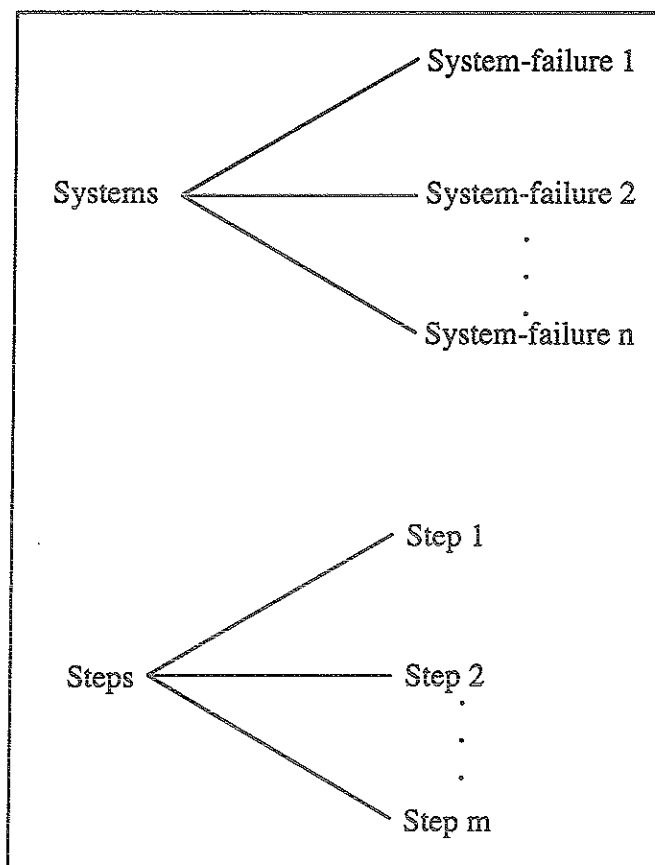


Fig. A3.5: Classification

In total, the hierarchy shown in Fig. A3.6 is thus obtained, and will be gradually extended by the user entering new system failures. The figure employs the names used in KEE. Since the slots in the class units "Systems", "Procedures" and "Steps" are defined, the lowest units in the

hierarchy are member units in order to ensure correct inheritance. These connections are represented as broken lines.

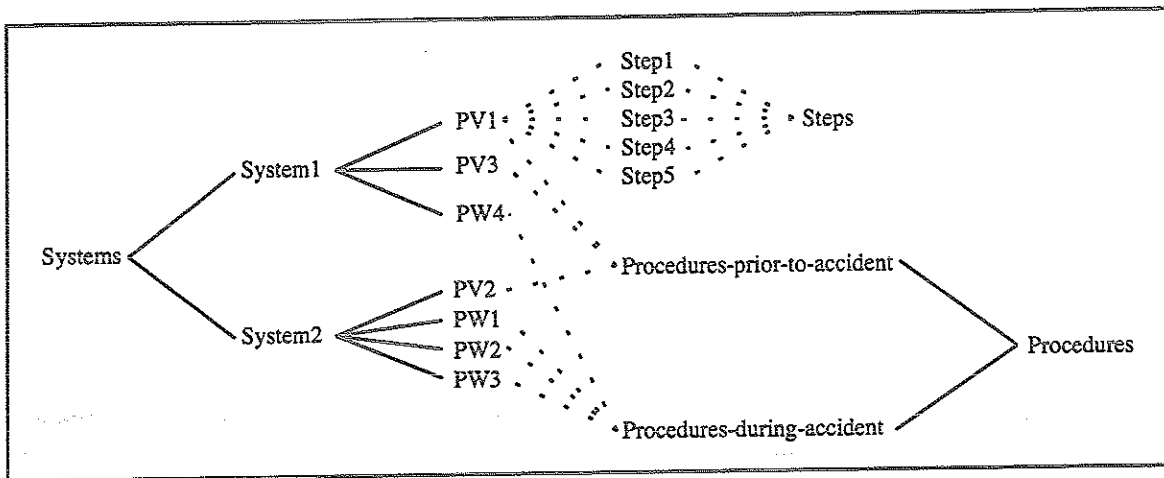


Fig. A3.6: Unit hierarchy of THERP in KEE

This modelling would also permit one procedure to be allocated to several system failures and one step to several procedures. However, this possibility has not yet been implemented.

If a new system failure, a procedure or a step is entered (cf. A3.3.2), the programme automatically generates the corresponding unit with the connections to the proper class units. Prior to the first input, the user therefore only has the hierarchy shown in Fig. A3.7. This hierarchy grows with the inputs.

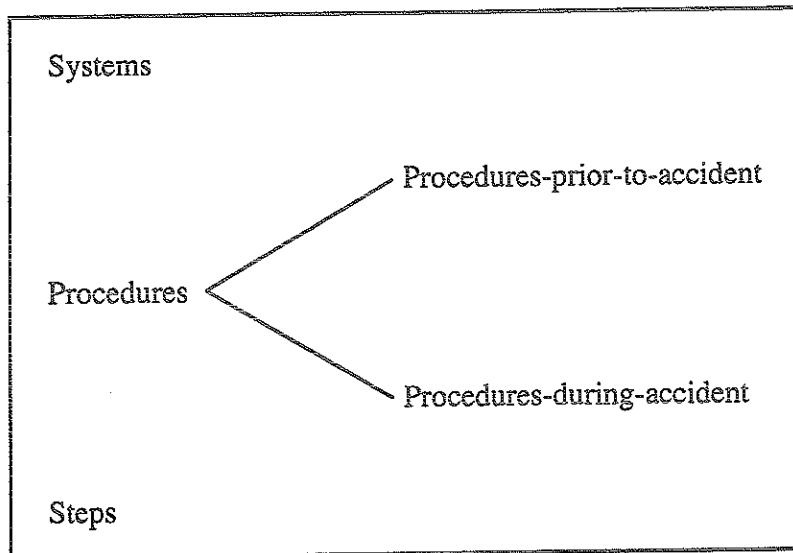


Fig. A3.7: Hierarchy prior to the 1st programme start

A3.3.1.2 Slots

The name of units generated automatically is composed of the higher-level class (e.g. system) and a number. In order to enable the user to later identify the system failure, procedure or step

to which the unit belongs, a short explanation is entered for every unit and stored in the "description" slot. In order to make this slot available in every unit, it is defined as a member slot in the "Systems" unit.

In performing a procedure it is important to observe the correct sequence of steps. Since it cannot be assumed that the user always specifies the steps in the correct order and since, moreover, subsequent extensions must be possible, this sequence is entered and stored in the "sequence" slot.

The next requirement is a number of slots in which the different groups of steps can be stored (e.g. set of all steps leading directly to system failure).

These slots are:

- "direct steps" all steps leading directly to system failure
- "indirect combinations" all combinations of steps leading indirectly to system failure
- "critical failure paths" all steps and combinations representing a unit, i.e. all failure paths to be included in the calculations

All slots containing groups of steps as well as the "sequence" slot are defined in the unit "Procedures" and are thus available in each procedure unit. Their value is always a list of step unit names, and combinations of steps are also represented as a list (cf. A3.3.2.3).

A number of questions must be answered for each procedure and each step. An answer slot is provided for every possible question, viz. slots "ans-1.1" to "ans-1.19" for steps, slots "ans-2.1" to "ans-2.7" for procedures prior to the accident and slots "ans-3.1" to "ans-3.8" for procedures during the accident. The possible answers to the questions form the value set for these slots. Slot "ans-2.1", for example, contains the answer to the question "the procedure is performed by an ..." and has thus the value range ("experienced operator", "inexperienced operator").

Before the calculations can be carried out, all data must be available. Instead of asking prior to each calculation whether all data have been entered, i.e. instead of checking the data slots for values in every unit belonging to the system failure hierarchy, a member slot "data-available" is defined in the "Systems" unit, which changes dynamically (active value) with the data input. In order to ascertain whether all data for the procedure have been entered, two additional slots defined in the "Procedures" unit are required, namely "data-questions" and "data-steps". The values and conditions with which these slot values change are compiled in Tab. A3.10.

All calculated values must be stored. For this purpose, the slots "BF" for basic failure probabilities, "MF" for modified failure probabilities, "KF" for correction failure probabilities, "FB" for failure conditions, "SF" for step failure probabilities and "FP" for failure path probabilities are defined in the unit "Steps" and the slot "PF" for procedure failure probabilities is defined in the unit "Procedures". Each of these slots contains a list with the corresponding probabilities; for example, slot "BF" contains the list (BF_{U1} , BF_{U2} , BF_{U3} , BF_{W1} , BF_{W2} , BF_{B1} , BF_{B2} , BF_S). Each probability is also a list containing all parameters for the probability in the form (median, uncertainty factor, mean, variance, upper bound, lower bound).

Unit-Typ	Slot	
Step	data-available	
	yes Questions are answered for this step or Step doesn't belong to an important failure path	no The step is newly entered
Procedure	data-questions	
	yes Questions are answered for this procedure	no The procedure is newly entered
Procedure	data-steps	
	yes Sequence end groups of steps are determined	no The procedure is newly entered or New steps are entered
Procedure	data-available	
	yes data-questions has the value 'yes' data-steps has the value 'yes' data-available has the value 'yes' in all steps belonging to the procedure	no The procedure is newly entered or One of the yes-conditions isn't fulfilled
System	data-available	
	yes data-available has the value 'yes' in all procedures belonging to the system-failure	no The system-failure is newly entered or One of the yes-conditions isn't fulfilled

Tab. A3.10: Dynamic data slot changes

A3.3.2 Inputs

The user must enter the hierarchy for each system failure, i.e. all procedures and their steps. He must tell the system which steps lead directly to system failure and which combinations of steps lead indirectly to system failure and which of these failure paths are used for further calculation. Moreover, a number of questions must be answered for each procedure and each step. The total input is controlled by a menu.

A3.3.2.1 Entering the hierarchy

A defined sequence must be observed in entering the hierarchy. A procedure can only be entered if it is known to which system failure it belongs. Moreover, the higher-level procedure must be defined for every new step. A system failure or a procedure is considered to be defined

as higher-level if it was either selected from the list of already existing system failures or procedures or if it is newly entered. The menu always permits only those operations which are possible at this time. Tab. A3.11 shows the different possibilities.

Current state	No system-failure defined; no procedure defined	System-failure defined; no procedure defined	System-failure defined; procedure defined
Possible unit definitions	Select system-failure Enter new system-failure	Select system-failure Enter new system-failure Select procedure Enter new procedure	Select system-failure Enter new system failure Select procedure Enter new procedure Enter steps

Tab. A3.11: Conditions for input

In Fig. A3.8 one will find an example of the menus for entering a new procedure for a system failure already existing in the hierarchy. This example is based on Fig. A3.6.

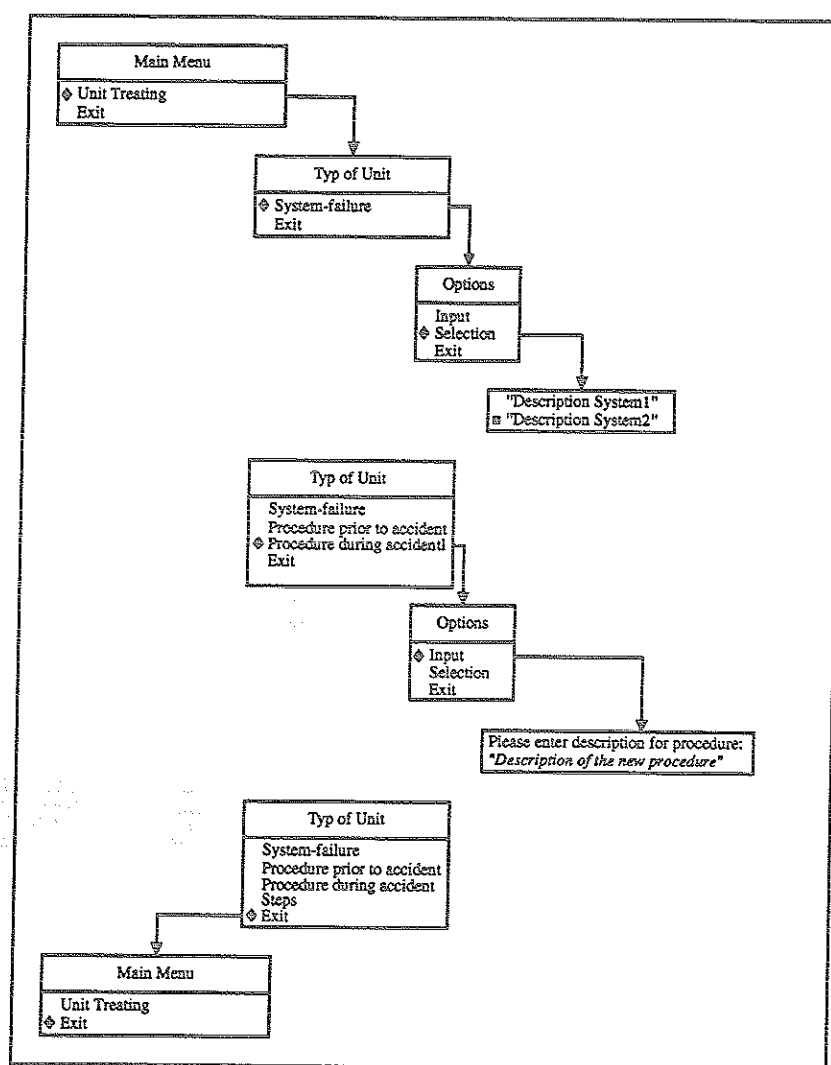


Fig. A3.8: Menu

After having executed these steps, a new "PW5" unit is generated which has the unit "System2" and the unit "Procedures prior to the accident" as the higher-level unit.

A3.3.2.2 Data input for procedures

A number of questions must be answered for each procedure. As soon as a procedure is selected using the menus, this data input is basically possible. The option "Data Input: Procedure" therefore appears additionally in the main menu. If the user starts the data input for a procedure, the programme first of all checks whether data have already been entered for this procedure ("data questions" slot has the value "yes"). If this is the case, the user is informed accordingly and can decide whether or not he will change these data.

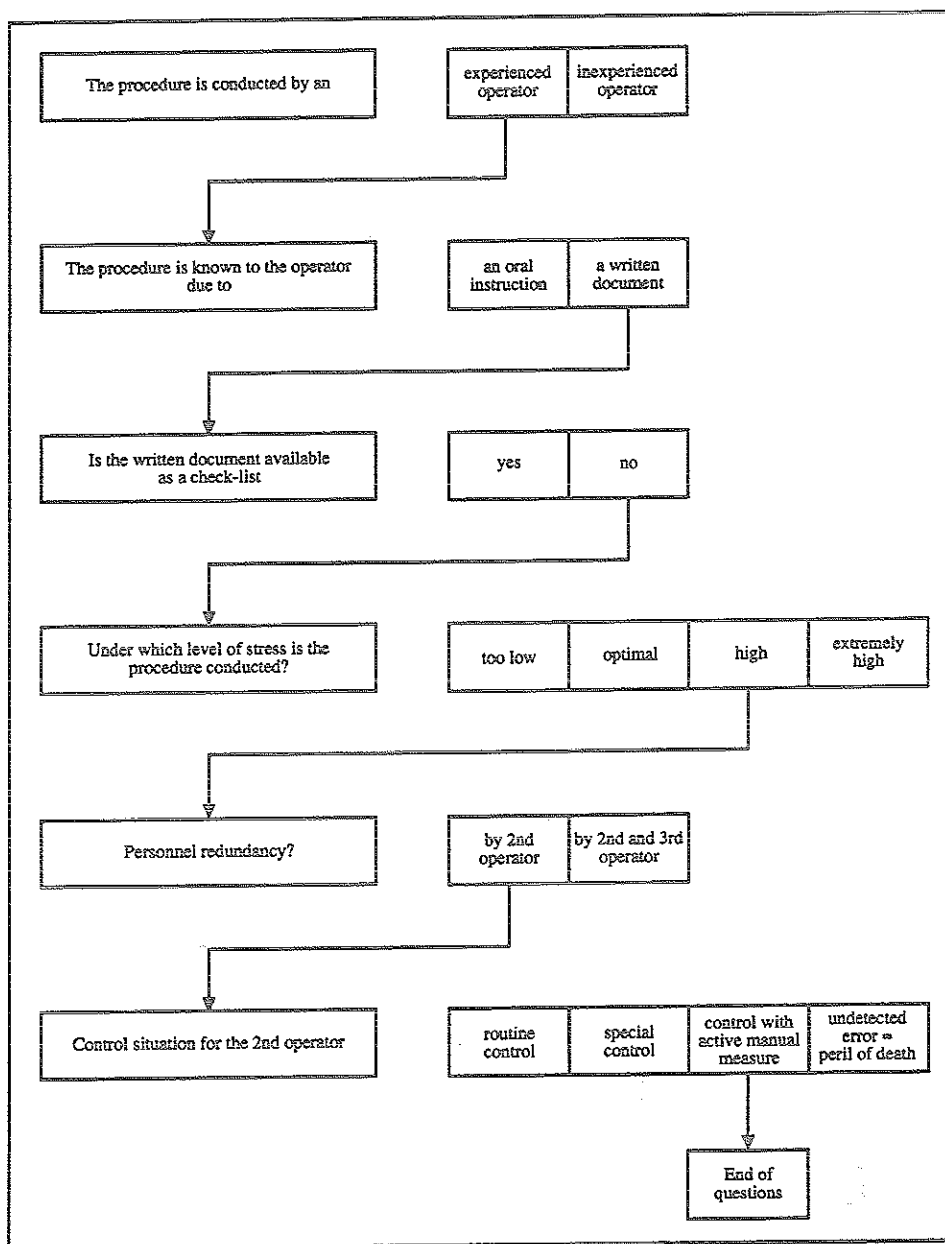


Fig. A3.9: Data input for a procedure prior to the accident

The actual data input for the procedure begins with the output of the first question and the associated possible answers. After having selected an answer (using the mouse), further questions and associated answers appear as a function of the first answer. Errors in data input can only be corrected by a completely new input. Fig. A3.9 shows the data input for a procedure prior to the accident (cf. also Tab. A3.1).

A3.3.2.3 Data for steps

If a procedure is defined and steps have been entered for this procedure, the data for these steps can be entered. The option "Data Input: Steps" appears in the main menu.

The data for the steps are entered in two sections. First of all, the step sequence must be entered and the steps of importance for further calculations (i.e. the step belongs to an important failure path) must be identified. The catalogue of questions must then be completed for each of these steps. Before starting the actual data input, the programme is asked whether data are already available. The different possibilities are compiled in Tab. A3.12.

Current state	Reaction of the programme
In the procedure-units: Slot 'data-steps' has the value 'no'	Enter sequence and grouping
Slot 'data-steps' has the value 'yes'	Question whether sequence and grouping should be changed
In the step-units: Slot 'data available' has the value 'no'	Answer questions for the step
Slot 'data available' has the value 'yes'	Question whether data for the steps should be changed

Tab. A3.12: Conditions for step data input

For entering the sequence, a list of the descriptions of all steps belonging to the procedure is printed out by the programme. These steps are consecutively marked using the mouse, so that a sequence is defined. After each marking the consecutive number before the step is shown so that the user has control of the sequence entered. Any numbered step is locked against further marking. The input is automatically terminated after marking the last step. Fig. A3.10 illustrates an input where the step sequence is entered for a procedure with 8 steps.

After having entered the sequence, the next input is the group classification of steps. First of all, all steps are identified which directly lead to system failure. For this purpose, the list of all steps is shown in the correct sequence. All steps leading directly to system failure are then marked using the mouse. All marked steps are deleted from the screen. The input end is indicated by marking the end line. Fig. A3.11 shows the input of direct steps, in this case steps 5 and 8.

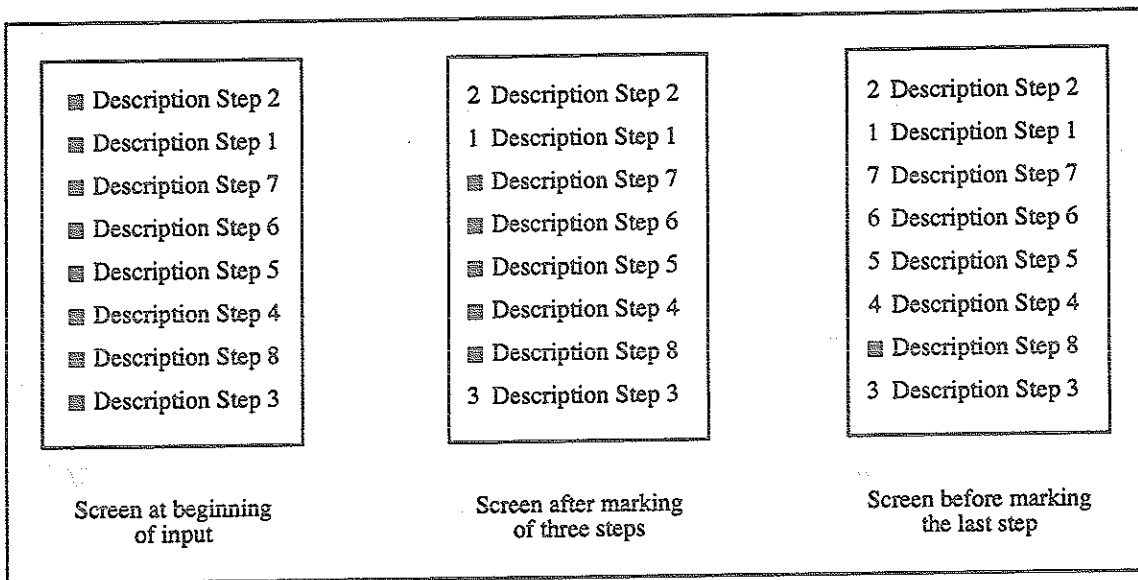


Fig. A3.10: Entering the sequence

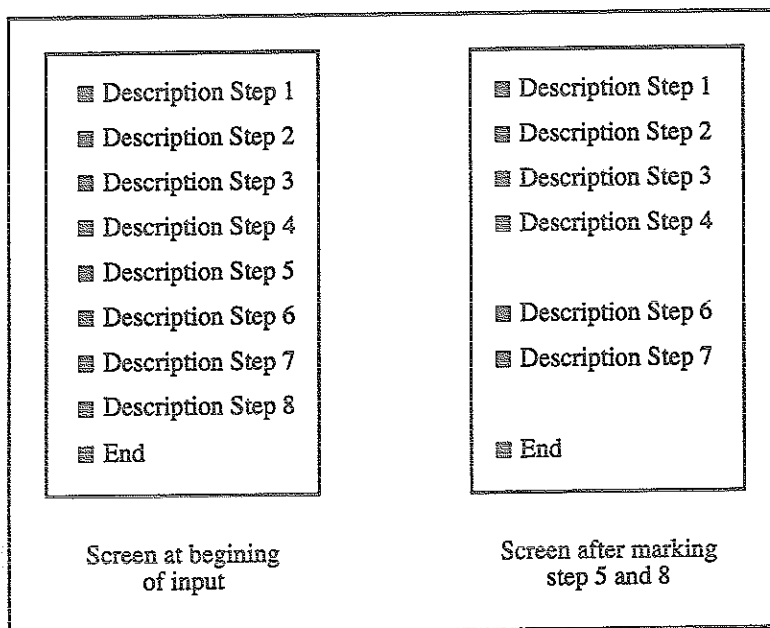


Fig. A3.11: Entering the direct steps

The next input is the determination of all combinations of steps which indirectly lead to system failure. The user is asked before entering a combination whether he wants to add another combination. A list of all steps not directly leading to system failure is shown for each combination. In this list, all steps are marked which belong to the combination to be entered. Marked steps are flagged with the number of the current combination and locked against further marking. Fig. A3.12 shows the input of the combinations from steps 1, 3 and 7 and from steps 1, 4 and 7.

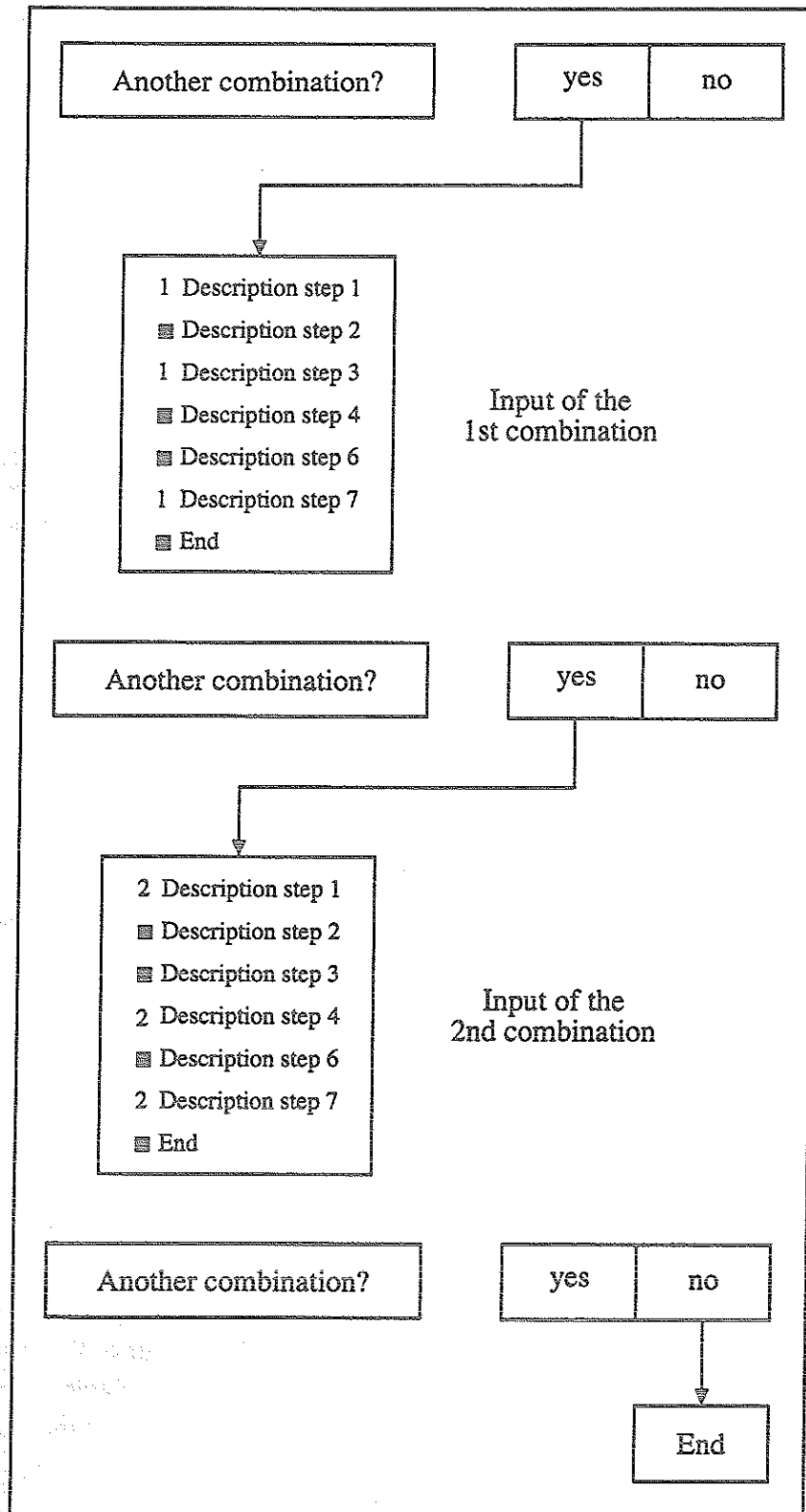


Fig. A3.12: Entering the two combinations

The last data input necessary in this section is the definition of the failure paths used for further calculations. In contrast to the description in 2.1, no units are formed selecting a representative failure path from these units, but the important failure paths are directly marked in the set of

failure paths. Our example in Fig. A3.13 shows the selection of failure path 1 consisting of step 5 and failure path 3 consisting of the combination of steps 1, 3 and 7.

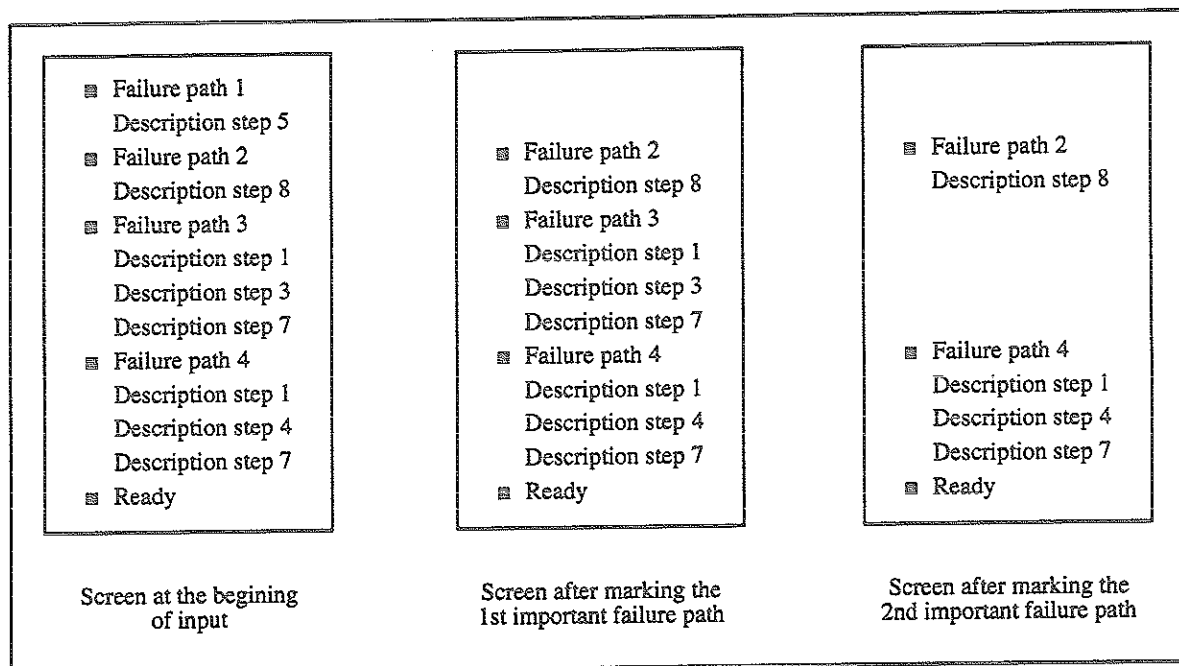


Fig. A3.13: Entering the important failure paths

If all important failure paths are marked, the catalogue of questions is completed for each step in this failure path. This input is the same as described for the data input procedures.

Since the data input for the steps (consisting of the sequence, group and data inputs for the individual steps) is started as a whole, input errors can only be corrected by restarting the complete data input for steps. However, the system checks for already existing data (see Tab. A3.12) so that individual steps in this overall input can be skipped.

A3.3.3 Calculation of the probability and output of results

The failure probability of a procedure can be calculated after complete data input. If the calculation is only performed for the procedure just selected, the value "yes" in the "data available" slot of the procedure unit is sufficient. If the probability for all procedures of the system failure just defined is to be calculated, the "data available" slot of the system unit must have the value "yes". Depending on the values in the "data available" slot, the option "Calculate current procedure" and/or "Calculate all procedures" appears in the main menu.

The median and uncertainty factor of basic failure probabilities, modified failure probabilities, correction failure probabilities and the failure conditions of all steps of a procedure are calculated using KEE rules. The rules are ordered in groups by a KEE hierarchy and the groups are started by a LISP programme. The hierarchy is shown in Fig. A3.14.

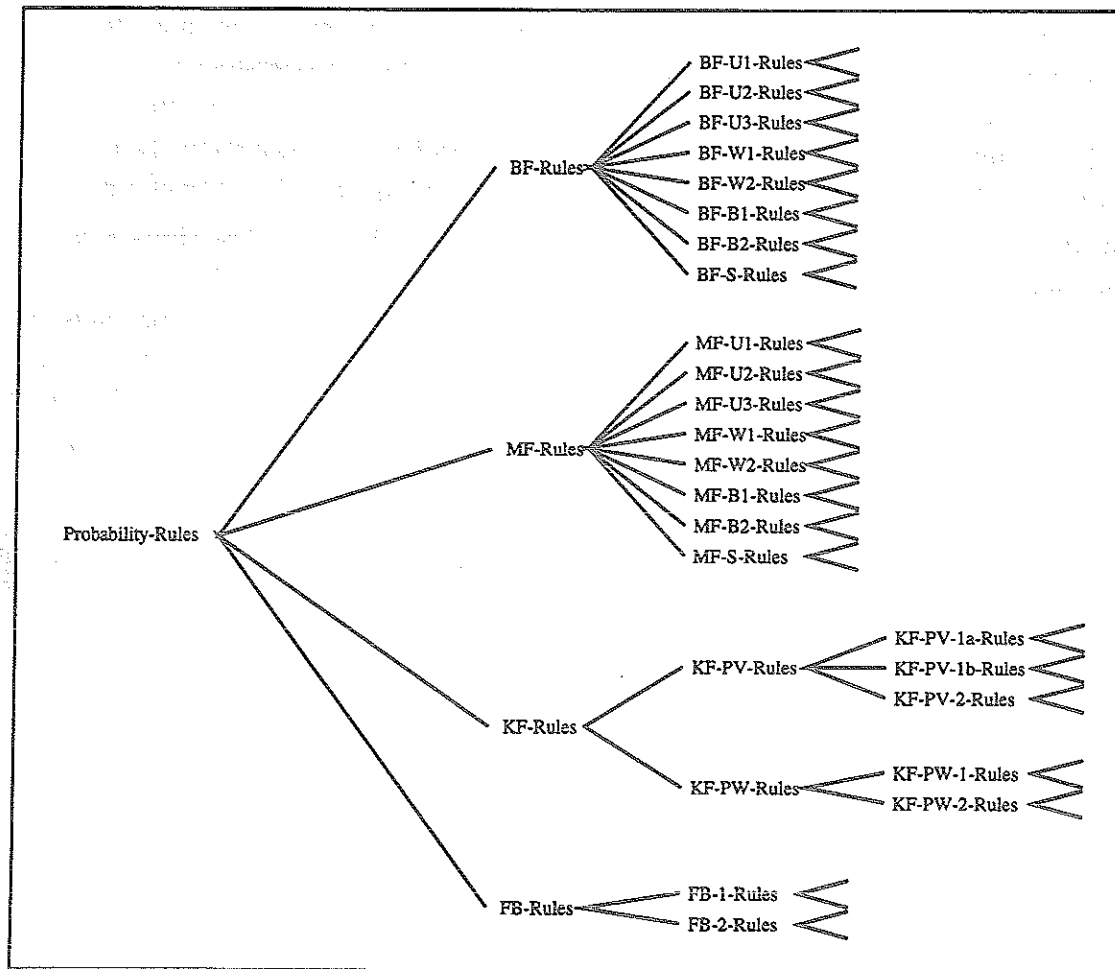


Fig. A3.14: Rule hierarchy

Each of these groups is the conversion of all conditions to be observed for calculating the corresponding probability. If these conditions include general features of the procedure (e.g. stress level of the operator), the group is again subdivided into a rule class for procedures prior to the accident and one for procedures during the accident. For example, the MF-W1 rule group (cf. also the conditions in Tab. A3.7) consists of the rule classes MF-WF1-PV rules and MF-W1-PW rules.

A rule in the MF-W1-PV rule class, for example, has the form:

IF

(?step belongs to the classes steps and *procedure*)
 (Slot ans-2.5 in *procedure* has the value "optimum")
 (Slot ans-2.1 in *procedure* has the value "inexperienced")
 (Slot ans-1.1 in ?step has the value "high")

THEN

(Modify the median and uncertainty factor of MF_{W1} in the MF slot of ?step)

Of course, this rule is not represented here in the formal KEE language, but in a more readable form. ?step is a variable, which consecutively receives as its value all the names of the units belonging to the class of steps and to the class *procedure* (name of the current procedure unit). This variable bond is achieved by the first condition in the IF part. The other conditions

ask for the values of individual slots, "ans-2.5" giving the answer to the question for the operator's stress level, "ans-2.1" the answer to the question for the operator's experience and "ans-1.1" the answer to the question for the cognitive level required for performing the current step. If all slots have the corresponding values, the median and uncertainty factor of the modified probability MF_{W1} are calculated for the current step using a LISP programme.

It can also be seen from Tab. A3.7 that one KEE rule is sufficient for a low stress level since the operator's experience and the cognitive level required for performing the step are not of interest in this case. Three rules are needed for optimum stress level since a distinction is made between experienced and inexperienced operators; for an inexperienced operator the cognitive level required for performing the step is divided into high or into medium and low. For high or extremely high stress levels there are four rules (subdivision into experienced or inexperienced operator and subdivision of the required cognitive level into high or into medium and low).

The MF-W1-PW rule class practically contains the same rules, the answer to the question of the operator's experience being in slot "ans-3.1" and the answer to the question of the operator's stress level being in slot "ans-3.6". The entire hierarchy for the MF-W1 rule class is shown in Fig. A3.15.

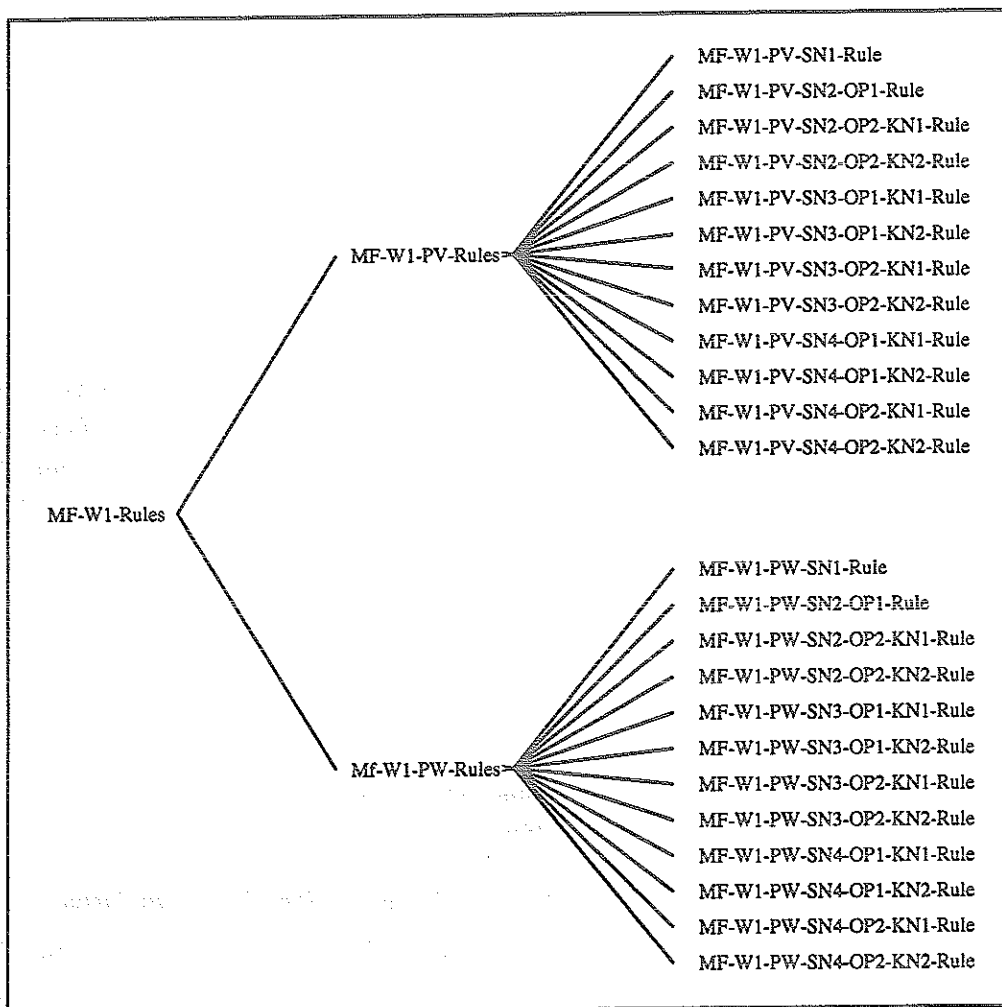


Fig. A3.15: Rule hierarchy for W_1

The KEE capability for processing chained rule systems is not needed for these calculations since only one rule is valid in each case. The use of a LISP programme for all calculations was not considered since it can be foreseen that new conditions will have to be inserted in the rules and new rules in the rule system for further developments. This poses fewer problems in a rule system than in a LISP programme, accepting the disadvantages that too many interrogations may be needed for processing the rules and that the calculations are performed for all steps and not only for the steps in important failure paths.

After processing all rules, the other parameters (lower bound, upper bound, mean and variance) of the probabilities determined by rules and all parameters of the probabilities of the error type contribution, the step failure probability, the failure path probability and the procedure failure probability are calculated by a LISP programme and/or active values.

Furthermore, the programme generates a data set containing the parameters of the procedure failure probability for all procedures which had to be calculated. This data set can be printed out on request.

In order to provide the user with a complete survey of all possibilities of the programme, all items of the main menu and the conditions for their appearance are compiled in Tab. A3.13.

Options	Is possible when:
Unit treating	Always
Data input: Procedure	The procedure is defined
Data input: Step	The procedure is defined and steps entered for this procedure
Calculate current procedure	The procedure is defined and the slot "data-available" for this procedure has the value "yes"
Calculate all procedures	The system failure is defined and the slot "data-available" for this system failure has the value "yes"

Tab. A3.13: Main menu

Appendix 4

Examples of expert systems supporting Probabilistic Safety Analyses

A4.1 LESSEPS

LESSEPS /Ma90/, developed at Electricité de France, is designed for the computerized management of complex reliability studies. The functions assumed to meet this objective are: reliability data management (basic elementary data, intermediate computation results, final results), management of the reliability models required for the various partial studies, communication of the data between the various computer codes used to sequence and quantify the models, and management of the sensitivity studies performed to assess the relative impact of the input parameters on the final results.

The LESSEPS software consists of the Common Data Structure (CDS). This structure serves as a communication interface between the computer codes used together for a given study. By merging several CDSs, a consistent grouping of the data relating to the corresponding studies was achieved.

Within the LESSEPS software two expert systems: EXPRESS and EXPGSI are used, developed in the framework of the PSA 1300 project to facilitate the study of complex systems in NPPs. These expert systems are applied to generate reliability models automatically. In the general context of the computerized study they are used before the LESSEPS software to generate the "parameter files" that will be managed by LESSEPS.

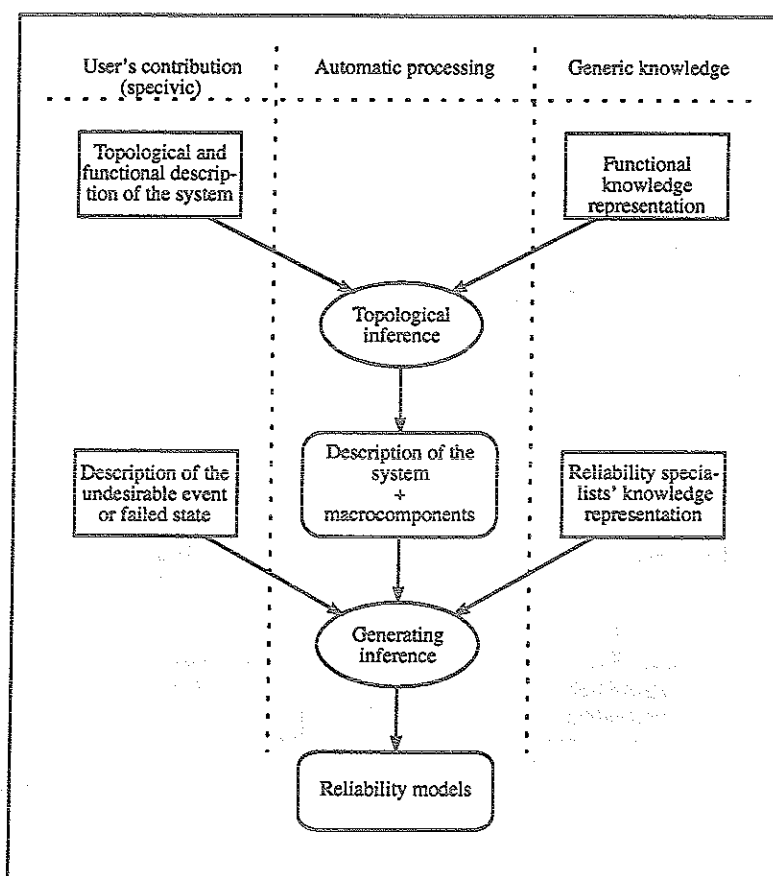


Fig. A4.1: General principles of EXPRESS and EXPGSI expert systems /Ma90/

The method revolves around two main inferences (Fig. A4.1). The first, or "topological inference", is aimed at grouping the components of the studied system into macrocomponents, according to the consequences of their main failures. The second, or "generating inference", is aimed at relating the component failure causes to the main failure modes and at deducing the failure consequences on the system in terms of "path" losses.

The systems studied belong to two main types: first, time-independent static systems and, second, time-dependent sequential systems. EXPRESS is used for static system analysis and generates a PHAMISS model of fault-trees. EXPGSI is used to analyse sequential systems and to generate a GSI model of states-graphs (Fig. A4.2). The topological inference rules used are the same for both expert systems but the selection of inference rules varies according to whether the study concerns a static system or a sequential system.

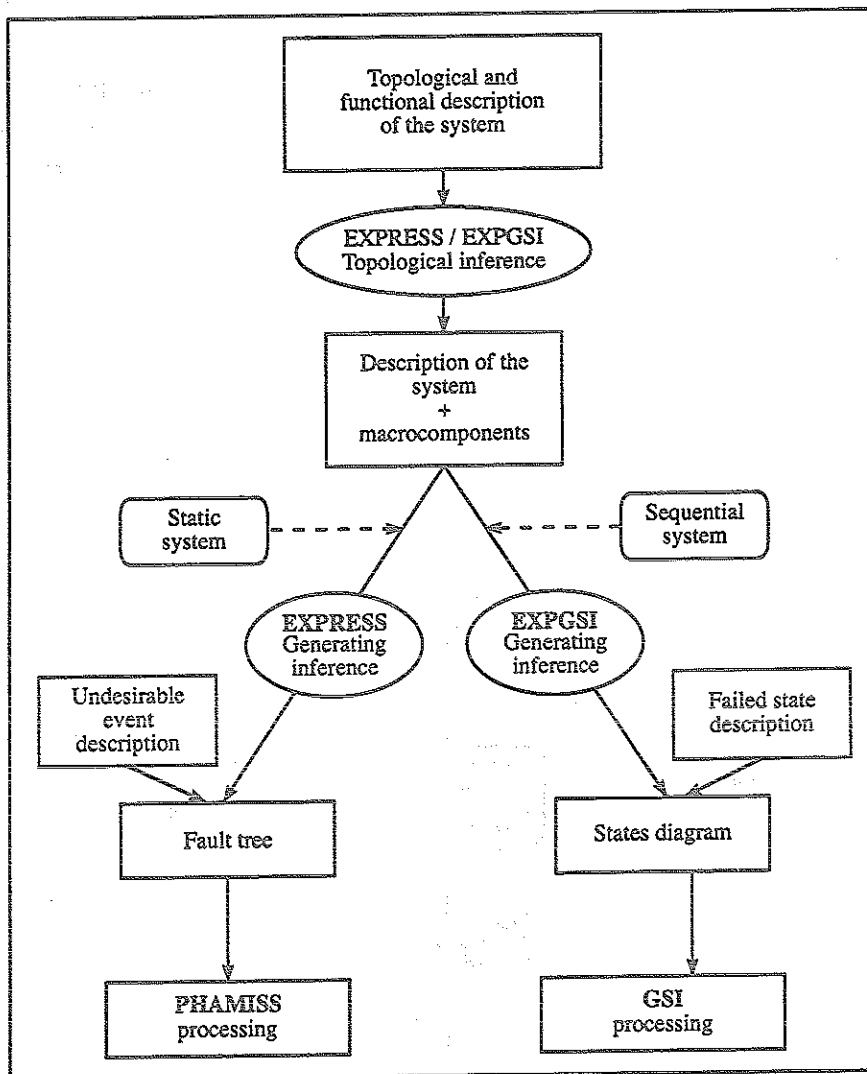


Fig. A4.2: EXPRESS and EXPGSI processing sequences /Ma90/

The general principle behind these expert systems stems from the two following general observations that most systems may be compared to sets of "lines" circulating a "fluid"

(coolant or electrical current) and that most of the failure modes of components in these systems can be grouped into two "main failure modes" according to their consequences: fluid blockage (open circuit for electrical energy) or outleakage (short circuit for electrical energy).

In the LESSEPS software a reliability study is based on the elementary data from statistical assessments of the operating experience including error factors. The overall uncertainty is evaluated by means of a Monte Carlo simulation.

The LESSEPS software product is run on an IBM-3090, but will be available on a SUN-type work station under the UNIX operating system. The data base management system used within the LESSERS software is the DB2 on IBM 3090 and ORACLE on SUN. The EXPRESS and EXPGSI expert systems are now run on IBM 3090. The inference engine used today with them is ALOUETTE (the first order inference engine developed by the Direction des Etudes et Recherches at EDF).

The work station for reliability engineers was designed as a part of a several-year long programme aimed at developing a software package. This software product is directly inspired by the existing products: LESSEPS, EXPRESS and EXPGSI. This work station should serve the following essential purposes /Ma90/:

- accumulate and centralize knowledge,
- structure and formalize knowledge,
- adapt the methods to the objective,
- manage reliability studies, and
- establish a dialogue between the designers, reliability engineers and operators.

A4.2 PRA-EXPERT and REX

PRA-EXPERT /Wa88B/ and REX /Wa90/ expert systems developed at the University of Maryland have been designed with an aim to perform PSA automatically. In order to provide a knowledge base for an AI programme, a single visual display of the plant reliability structure was developed, called the Master Plant Logic Diagram (MPLD) /Wa88B/. This diagram shows all the interrelationships amongst initiating events, front-line systems, and support systems. Fig. A4.3 shows a very simple MPLD. The MPLD also shows the topology of systems and subsystems in some detail.

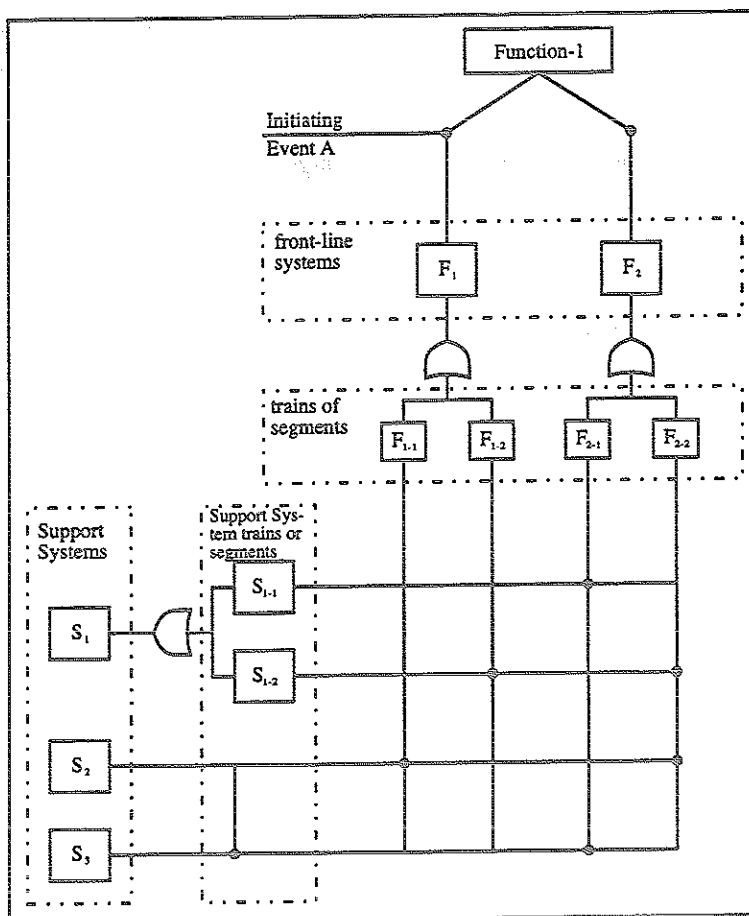


Fig. A4.3: A simple Master Plant Logic Diagram (MPLD) /Wa90/

The success-oriented MPLD determines the functional requirements which must be satisfied if a core damage event is to be prevented, and then by identifying the front-line systems which have the capability to support the functional requirements. Each function is then divided into the front-line systems which have the capacity to individually or collectively support these functional requirements (e.g. **F₁** and **F₂** in Fig. A4.3). The relationships between initiating events and front-line system requirements are logically described in terms of functional success criteria, whereby each required function is satisfied if e.g. either: no LOCA or transient is present, or the system hardware is available.

The hierarchy of the MPLD model is further displayed by the support system matrix. Each front-line system is divided into its independent segments (e.g. trains or subtrains, **F_{1,1}** ...,

etc.), and the effect of failures from all combinations of support systems or their segments (S_1 , S_2 and S_{1-1} , S_{1-2} , ..., etc.) are established and explicitly shown on the crossing of lines by '•' (a bold dot). An analogous process is followed to establish interrelationships between different support systems. The MPLD model provides an opportunity to eliminate the need for analyst modelling of the support system event tree.

In this expert system also other intelligent activities such as data evaluation and analysis, human recovery actions, and event tree construction have been considered. The main objective to develop an expert system called Reliability EXpert (REX), being a part of the PRA-EXPERT, was to develop a tool to deal in an intelligent way with the topology of systems, e.g. the system diagram. This required the use of graphical representation of systems in the expert system, and direct recognition of these graphs represented as objects by the computer systems.

REX is an intelligent decision and analysis aid for reliability and risk studies with focusing on the graphical representation of systems using intelligent graphics objects. An object-oriented programming approach is used to graphically and descriptively represent a complex system hierarchically in the computer /Wa90/. In this approach any item (component, block, connection, ..., etc.) in this system will be treated as an object. A collection of similar objects is called a class. It is possible to combine two or more classes into a new class.

As the basis for PRA logical modelling, the so-called 'large event tree - small fault tree' approach was selected, because of its modularity and the use of a support system state concept to create the ability to independently evaluate plant systems (both front-line and support systems). This approach offers a breakdown of the overall PRA model into several independent submodels (e.g. fault trees, plant states and event trees).

By using a set of rules the minimal success paths of the system analysed are determined. Then the probability of the system is evaluated. Because the success paths are not disjoint, in order to minimize the effect of inclusive items among the minimal success paths, the method of sum of disjoint products is applied /Wa90/.

The computer hardware used for building REX is LMI-Lambda 2 x 2TM Lisp machine. The entire programme is written using ZETALISP-PLUS version 3. Recently, most efforts undertaken have been to develop a PC version of PRA-EXPERT. Future plans include the development of a more advanced version of REX on both PCs and a SUN-360 work station using G-2 (G-2 is a trademark of Gensysm Inc.) object-oriented expert system shell /Wa90/.

A4.3 SAPIEXT

SAPIEXT /Me90/ aims to provide an expert system environment which integrates some of the computer programmes currently employed at SRD, UK for non-nuclear PSAs. The main advantages of such a system include:

- (a) retention of expert knowledge,
- (b) improvement in the efficiency and standard of PSA's,
- (c) rapid training aid.

In addition the advantages of an integrated expert system approach will be to ensure that the programmes are accessed in a consistent manner and that much expert help will be given to the user. This includes the use of rule tracing, which will allow the user to examine the logic behind a decision which the system has made.

A series of prototypes have been developed, using several expert system packages, to demonstrate certain features of SAPIEXT and determine whether these packages were capable of meeting the project requirements. These prototypes have been developed using: the rule-based expert system shell Xi Plus, an expert system toolkit GoldWorks II and an expert system environment which combines hypertext and expert system technology KnowledgePro.

After assessing the prototypes, Xi Plus was selected as the development for SAPIEXT because it was PC-based and enabled external programmes to be called directly from a knowledge base. Xi Plus is a rule-based shell for both building and consulting expert system applications. An application can consist of one or more knowledge bases, each made up of knowledge represented as identifiers, facts, rules, questions, defaults or demons.

At the heart of SAPIEXT will be the core expert system controlling the consultation, guiding the user through the stages of a safety assessment and pin-pointing the appropriate assessment code. Once the correct code has been identified all the information required to run that programme will be extracted from the user. Where necessary, an input file will be created in the format required by the programme. The programme would be then run directly from the expert system and the results displayed to the user for further analysis.

The primary functional and operational requirements for SAPIEXT were assumed to include functions typical of expert systems developed in multi-data and programmes environment and those associated with Quality Assurance (QA) purposes /Me90/. It is believed that by developing knowledge-based systems, such as SAPIEXT, to perform the routine part of an expert's job, the efficiency of the expert can be enhanced, thereby allowing more time to be spent on non-routine tasks.

A4.4 STARS

STARS /Po90/ is a knowledge-based Software Tool for the Analysis of Reliability and Safety, being developed at the CEC Joint Research Centre, Ispra, Italy. The overall aim of the STARS project was to develop computer support for the various tasks involved in safety and reliability analysis of complex plants. Among these are the identification of potential hazards, the identification of events and event sequences leading to hazards, the construction of event trees, fault trees and Markov models, the logical and/or probabilistic analysis of the models and the identification and treatment of common mode and common cause failures.

The functions covered by STARS can be seen from the overview scheme presented in Fig. A4.4. They include:

- (1) Initial data collection and updating of the information: this includes the description of the plant or system under analysis in terms of the functions, the processes involved, the P&I diagrams and functional diagrams of systems, the characteristics of the components and the description of the system operation, process control and protection.
- (2) Qualitative analysis: for the identification of hazardous events or event sequences and ranking of such events and event sequences in terms of the severity of their consequences. This task can be performed by Hazard and Operability Analysis (HAZOP) and/or Failure Mode and Effect Analysis (FMEA).
- (3) Event sequence modelling: for the structuring of the identified events and event sequences into logic models. This includes construction of event trees and/or master fault trees.
- (4) Systems modelling: for the construction of logic models for the system malfunctions that appear in the event sequences or event sequence models. This includes the construction of fault trees and the generation of states graphs or transition matrices. Moreover, modelling includes the identification of dependencies between subsystems or components.
- (5) Model analysis: for the logic and /or probabilistic analysis of the constructed models for systems and event sequences. This includes event tree analysis, fault tree analysis, Markov analysis and possibly reliability simulation. Moreover, the analysis phase also considers the quantification of dependencies if dependent events have been identified in the previous task.

In order to provide the user with 'intelligent support', a knowledge-based approach was adopted for realizing the qualitative analysis tools, the logic and probabilistic modelling tools and the dependency identification tool. These tools will draw their knowledge from four knowledge bases (KBs): a plant/unit KB, a substance KB, a reaction matrix KB and a component KB. The first three are used for performing HAZOP and consequence ranking. The last one is used for logic model construction. For the KB management and the efficient use of the rules in the KB, a powerful KB management system was developed. It is used by the inference engine to rapidly access the relevant parts in the KB.

The STARS project is based on the use of a UNIX work station. A small CAD tool has been implemented for graphical representation of information in the form of the system topology. For displaying, editing and handling the constructed fault trees, a powerful fault tree editor was developed. The fault tree part of STARS contains a programme for logical and probabilistic analysis of fault trees, based on an improved version of the SALP-PC software.

STARS is primarily for assessing plant safety and reliability but its capability can be considered in a more comprehensive way. Some possible and new types of exploitation are: use as 'real time' decision supporting tools for reliability and safety optimization under different design

alternatives and/or maintenance policies and use as a supporting tool for a living PSA management of specific plants /Po90/.

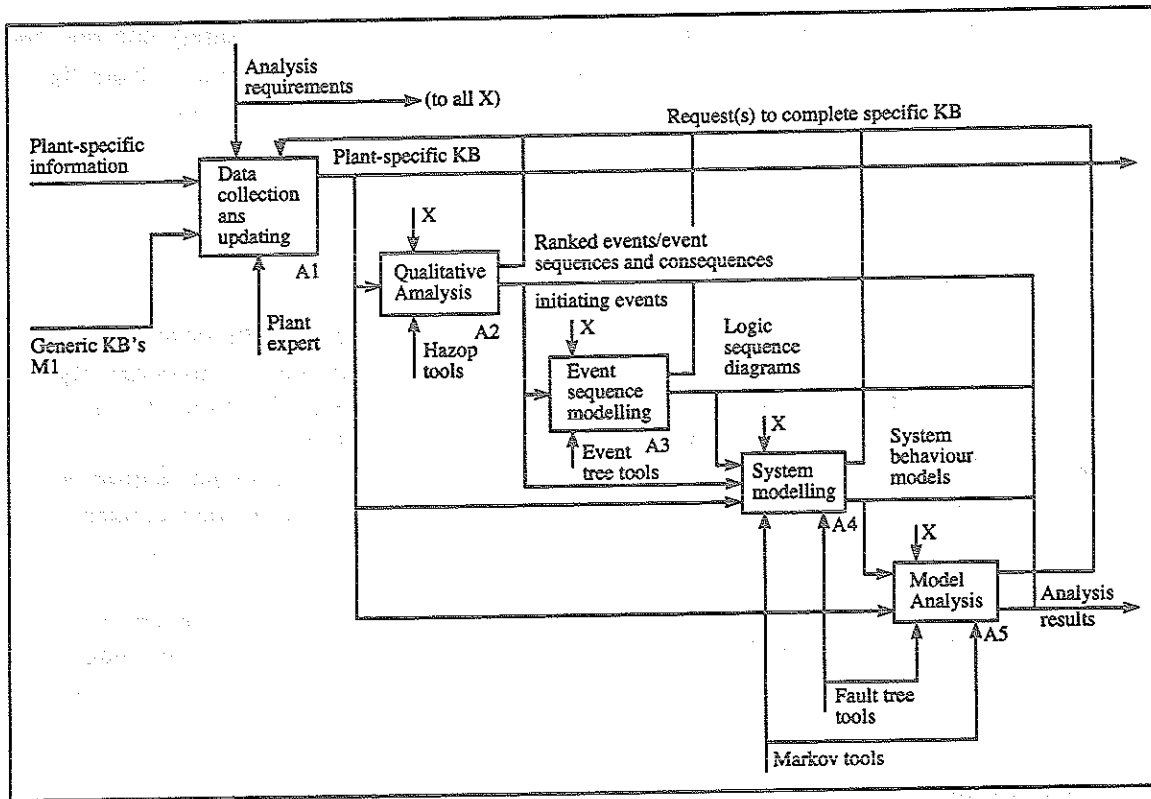


Fig A4.4: Overview of tasks covered in STARS /Po90/

A4.5 Other expert systems

ERNEST - Expert Reliability Net for Engineered System Treatment /Ga87a/, developed at the Politecnico di Milano, is a rule-based expert system for the interactive knowledge-driven construction of fault trees. The system is organized according to a number of knowledge-based modules that allow rules to be established, contain metarules, and collect and interpret data. Metarules, frames and semantic nets allow the evidence that may come from different sources to be combined. Computer graphics and first-order propositional logic are chosen as the means for structuring information.

To represent component states, the system models elementary components in terms of multiple-valued logic (MVL) mini-trees. These mini-trees are then combined into the computational MVL tree, which is the fault tree in its multiple-value logical form. Such trees can be reduced to conventional two-states fault trees. The system is claimed to be flexible and effective. Flexibility is achieved by its rule-based knowledge representation enabling easy modification of the existing tree. Effectiveness is claimed to be achieved by means of fuzzy algebra and related measures of uncertainty. The search for prime implicants and ranking heuristics under fuzzy logic and approximate reasoning was a part of the work still under way /Ga87a/.

HAZOPEX - HAZard and OPerability EXpert System /Ka87/, developed at the Technical Research Centre of Finland, is an expert system intended to support process designers and engineers in carrying out hazard and operability (HAZOP) studies of process systems, which is a method often used in the chemical industry. HAZOPEX takes the process and instrumentation diagram (P&ID) and complementary information on the chemical process and its components as input. It is designed to support HAZOP analyses of process systems in two ways. First, it identifies the potential causes and consequences of plant states that deviate from normal. Second, it provides suggestions for actions taken to prevent the plant deviation or to minimize its consequences. In order to restrict the search to a reasonable group of rules and manageable time, the knowledge representation method chosen for HAZOPEX is based on shallow knowledge and heuristic rules instead of deep knowledge. The HAZOPEX was designed for performing the qualitative analysis. It was not capable of carrying out quantitative analysis. Human errors were also not treated in detail. The expert system is developed with a LISP machine in a KEE environment.

PLEXSYS - PLant EXpert SYstem /Na89, Na90/ is specialized artificial intelligence software developed for expert systems involving complex systems such as electric power plants. The central facility in PLEXSYS is a model editor which enables users to build or represent their plant in schematic format similar to computer-aided design (CAD) systems. In addition to the schematics, there are data or knowledge base structures and methods which automate reasoning and problem solving tasks involving complex systems. An effort was undertaken was to automate the building of the PLEXSYS knowledge base directly from a CAD data base.

GDS - Generic Diagnostic System /Ne87/ is a general tool (framework system) which provides a skeletal software system to facilitate the utility personnel building a 'problem specific' diagnostic expert system for complex plant and process systems. Application examples include process monitoring and operational diagnosis, power plant operations and plant emergency response systems, hardware system diagnostics and maintenance advisors. The system

diagnoses malfunctions at two level: a functional level, and a root cause level. It is capable of handling multiple failures in complex situations. Because the system's primary logic model is a fault tree, a module for automating fault tree constructions has been included to facilitate building applications. The inferencing techniques include an enhanced forward-chaining algorithm and a collection of logic reasoning rules. A prototype system is LISP-based and runs on Motorola 68000 work stations and the VAX family of processors.

JUL-2928

June 1994

ISSN 0944-2952