NUREG/CR-4862
Vol. 1

Cognitive Environment Simulation: An Artificial Intelligence System for Human Performance Assessment

Summary and Overview

Prepared by D. D. Woods, E. M. Roth, H. Pople, Jr.

Westinghouse Electric Corporation

Prepared for
U.S. Nuclear Regulatory
Commission

NOTICE

Availability of Reference Materials Cited in NRC Publications

Most documents cited in NRC publications will be available from one of the following sources:

1. The NRC Public Document Room, 1717 H Street, N.W.
   Washington, DC 20555

2. The Superintendent of Documents, U.S. Government Printing Office, Post Office Box 37082, Washington, DC 20013-7082

3. The National Technical Information Service, Springfield, VA 22161

Although the listing that follows represents the majority of documents cited in NRC publications, it is not intended to be exhaustive.

Referenced documents available for inspection and copying for a fee from the NRC Public Document Room include NRC correspondence and internal NRC memoranda; NRC Office of Inspection and Enforcement bulletins, circulars, information notices, inspection and investigation notices; Licensee Event Reports; vendor reports and correspondence; Commission papers; and applicant and licensee documents and correspondence.

The following documents in the NUREG series are available for purchase from the GPO Sales Program: formal NRC staff and contractor reports, NRC-sponsored conference proceedings, and NRC booklets and brochures. Also available are Regulatory Guides, NRC regulations in the *Code of Federal Regulations,* and *Nuclear Regulatory Commission Issuances.*

Documents available from the National Technical Information Service include NUREG series reports and technical reports prepared by other federal agencies and reports prepared by the Atomic Energy Commission, forerunner agency to the Nuclear Regulatory Commission.

Documents available from public and special technical libraries include all open literature items, such as books, journal and periodical articles, and transactions. *Federal Register* notices, federal and state legislation, and congressional reports can usually be obtained from these libraries.

Documents such as theses, dissertations, foreign reports and translations, and non-NRC conference proceedings are available for purchase from the organization sponsoring the publication cited.

Single copies of NRC draft reports are available free, to the extent of supply, upon written request to the Division of Information Support Services, Distribution Section, U.S. Nuclear Regulatory Commission, Washington, DC 20555.

Copies of industry codes and standards used in a substantive manner in the NRC regulatory process are maintained at the NRC Library, 7920 Norfolk Avenue, Bethesda, Maryland, and are available there for reference use by the public. Codes and standards are usually copyrighted and may be purchased from the originating organization or, if they are American National Standards, from the American National Standards Institute, 1430 Broadway, New York, NY 10018.

NUREG/CR-4862
Vol. 1
RX

Cognitive Environment Simulation: An Artificial Intelligence System for Human Performance Assessment

Summary and Overview

Manuscript Completed: September 1987
Date Published: November 1987

Prepared by
D. D. Woods, E. M. Roth, H. Pople, Jr.*

With Contributions by
D. Embrey
Human Reliability Associates
School Lane, Dalton, England

Westinghouse Electric Corporation
Research and Development Center
1310 Beulah Road
Pittsburgh, PA 15235

*University of Pittsburgh and Seer Systems
Pittsburgh, PA 15235

Prepared for
Division of Reactor and Plant Systems
Office of Nuclear Regulatory Research
U.S. Nuclear Regulatory Commission
Washington, DC 20555
NRC FIN D1167

Abstract

This report documents the results of Phase II of a three phase research program to develop and validate improved methods to model the cognitive behavior of nuclear power plant (NPP) personnel. In Phase II a dynamic simulation capability for modeling how people form intentions to act in NPP emergency situations was developed based on techniques from artificial intelligence. This modeling tool, Cognitive Environment Simulation or CES, simulates the cognitive processes that determine situation assessment and intention formation. It can be used to investigate analytically what situations and factors lead to intention failures, what actions follow from intention failures (e.g., errors of omission, errors of commission, common mode errors), the ability to recover from errors or additional machine failures, and the effects of changes in the NPP person-machine system.

The Cognitive Reliability Assessment Technique (or CREATE) was also developed in Phase II to specify how CES can be used to enhance the measurement of the human contribution to risk in probabilistic risk assessment (PRA) studies.

The results are reported in three self-contained volumes that describe the research from different perspectives. Volume 1 provides an overview of both CES and CREATE. Volume 2 gives a detailed description of the structure and content of the CES modeling environment and is intended for those who want to know how CES models successful and erroneous intention formation. Volume 3 describes the CREATE methodology for using CES to provide enhanced human reliability estimates. Volume 3 is intended for those who are interested in how the modeling capabilities of CES can be utilized in human reliability assessment and PRA.

Table of Contents

List of Figures

Acknowledgments

The authors would like to express their appreciation to Thomas G. Ryan, the NRC project officer, for the guidance and support he provided throughout the model development program.

We would also like to thank Michael Coombs, Allen Newell, Richard Pew, Jon Young, and John Wreathall, who participated in an NRC sponsored workshop to review the model development work, for their helpful comments.

We are grateful to the many colleagues who discussed with us how a model of intention formation could aid HRA and PRA, especially V. De Keyser, O. Svenson, J. Rasmussen, B. Kirwan.

Special thanks are due to James Easter, Glenn Elias and Michael Mauldin for helping to keep the model development on-track to capture the nuclear power plant as a problem solving world.

1. Introduction

This report documents the results of Phase II of a three phase research program sponsored by the U. S. Nuclear Regulatory Commission to develop and validate improved methods to model the cognitive behavior of nuclear power plant (NPP) personnel during emergency operations. In Phase II a model of how people form intentions to act in NPP emergency situations (Cognitive Environment Simulation or CES) was developed using artificial intelligence (AI) techniques. A methodology for using the model to enhance measurement of the human contribution to risk in probabilistic risk assessment (PRA) studies (Cognitive Reliability Assessment Technique or CREATE) was also developed.

The results of the Phase II research are reported in three volumes. Each volume is a self-contained report on one perspective of the Phase II model development work. This volume provides an overview of CES and CREATE. Volume 2 of this report gives a detailed description of the structure and content of the CES cognitive model. Volume 2 is intended for those who want to know about how CES models successful and erroneous human intention formation. Volume 3 describes the CREATE methodology. It outlines the steps involved in using CES as part of human reliability analysis (HRA) in PRA studies, and it describes how CES can be used to better estimate human reliability. Volume 3 is intended for those who are interested in how the modeling capabilities of CES can be utilized in HRA and PRA.

1.1 The Importance of Modeling Operator Cognitive Activity for Human Reliability Assessment

The quality of human performance has been shown to be a substantial contributor to nuclear power plant safety. Some PRA studies have found that approximately one half of the public risk from reactor accidents can be related to human error (Levine and Rasmussen, 1984; Joksimovich, 1984). Studies of NPP operation and maintenance indicate from 30% to 80% of actual incidents in nuclear power plants involve significant human contribution (Trager, 1985). The analytical and empirical records clearly show that the human contribution to total safety system performance is at least as large as that of hardware reliability (Joksimovich, 1984).

A significant factor in determining human action under emergency conditions is *intention formation* -- deciding on what actions to perform.[1] Errors of

[1] This is contrasted with *execution of intentions* -- carrying out the sequence of actions decided upon.

intention are an important element of overall human contribution to risk, and the PRA community has recognized the need for more effective ways to capture this component of human error (Levine and Rasmussen, 1984).

The U.S. Nuclear Regulatory Commission has embarked upon a program of research to build a computer model of human intention formation (how people decide on what actions are appropriate in a particular situation) in order to better predict and reduce the human contribution to risk in NPPs. The model simulates likely human responses and failure modes under different accident conditions, comparable to the analytic tools available for modeling physical processes in the plant.

This research program consists of three phases. Phase I (completed in April of 1986) was a feasibility study which determined that it is practical to build such a cognitive model to provide useful input to human reliability analysis and probabilistic risk assessment (the results of the assessment are reported in NUREG/CR-4532). The feasibility study identified a specific modeling approach based on extensions and elaborations of an artificial intelligence (AI) problem solving system created by Dr. H. Pople, Jr. of the University of Pittsburgh and Seer Systems, Inc. for internal medicine applications (Pople, 1982; 1985)

Phase II of the research project focused on model development and application to HRA based on the approach identified in Phase I. Specifically:

1. A model of how people form intentions to act in emergency operations in NPPs was developed using AI techniques. The model, called Cognitive Environment Simulation or CES, is the first analytic computer simulation tool which can be used to explore human intention formation in the same way that reactor codes are used to model thermodynamic processes in the plant.

2. A methodology, called Cognitive Reliability Assessment Technique or CREATE, was developed which specifies how this capability can be used to enhance measurement of the human contribution to risk in PRA studies.

An additional phase of the research project is planned whose objective is to conduct field evaluation and validation of the CES cognitive model and the CREATE methodology.

1.2 The Role of Modeling of Human Intention Failures in Risk Analysis

Model development addressed one part of human behavior: human intention formation (deciding what to do) and erroneous intentions to act. This scope was chosen, first, because models and techniques are already available to assess the form and likelihood of execution errors in human reliability studies (e.g., Reason & Mycielska, 1982; Swain & Guttman, 1983). A second reason for selecting this scope is because erroneous intentions are a potent source of human related common mode failures which can have a profound impact on risk -- as actual accidents such as Three Mile Island and Chernobyl have amply demonstrated. Intentions to act are formed based on reasoning processes. The scientific disciplines that study these processes are called cognitive sciences or mind sciences and include a variety of fields such as cognitive psychology and artificial intelligence. Models of these processes are called "cognitive models."

In Phase II a computer simulation of intention formation in emergency operations was developed. This system, Cognitive Environment Simulation or CES, is the first analytic computer simulation tool that can be used to model human intention formation in the same way that reactor codes are used to model thermodynamic processes in the plant.

CES is a simulation of cognitive processes that 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, person-machine interfaces). By simulating the cognitive processes that determine situation assessment and intention formation, it provides the capability to establish analytically how people are likely to respond, given that these situations arise. This means one can investigate

- what situations and factors lead to intention failures,
- the form of the intention failure,
- the consequences of an intention failure including,
  - what actions will not be attempted -- errors of omission,
  - what actions the intention failure will lead to -- *commission errors* and *common mode failures*, that is, those leading to the failure of otherwise redundant and diverse systems due to misperception of plant state or another cognitive processing breakdown,

- *error recovery* -- whether the human intention failures or execution errors or failures of plant equipment to respond as demanded will be caught and recovery action taken (and information on the time until recovery),

- "improvised" action sequences (responses other than those specified in available procedures) that operators may take in different circumstances.

The ability of CES to predict errors of commission is particularly important since misapprehension of plant state by the operator can result in multiple actions which can have broad systemic effects. Intention failures are a major source of *human related common mode failures* -- multiple failures that are attributable to a common element (namely, the erroneous intention). Examples of this are cases where the situation is misperceived, and the operator deliberately decides it is appropriate to turn off multiple, otherwise redundant and diverse systems as occurred at Three Mile Island and Chernobyl. The PRA community generally recognizes the importance of identifying common mode failure points because they can have large and widespread effects on risk.

Because CES models the processes by which intentions to act are formed, it can be used, not only to find intention error prone points, but also to identify the sources of cognitive processing breakdowns and intention failures. This means that it can help to develop or evaluate error reduction strategies.

CES also provides an analytic tool for investigating the effects of changes in NPP person-machine systems including new instrumentation, computer-based displays, operator decision aids, procedure changes, training, multi-person or multi-facility (e.g., technical support center) problem solving styles. This means that proposed changes/enhancements to NPP person-machine systems can be analytically evaluated before they have been implemented.

CES, as a modeling environment, is a specific instance of an artificial intelligence problem solving system, EAGOL.[2] The EAGOL problem solving architecture embodies unique capabilities for reasoning in dynamic situations that include the possibility of multiple faults. CES uses these capabilities to capture the kinds of cognitive processes that contribute to intention formation.

[2]EAGOL is a software system and proprietary product of Seer Systems. EAGOL builds on the conceptual framework of the CADUCEUS AI problem-solving system developed for medical problem-solving applications (Pople, 1985).

Cognitive Reliability Assessment Technique (CREATE) is the method for using the capabilities of CES to better evaluate the potential for significant human errors in PRA analysis. 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 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 will arise are identified, a quantification procedure is used to assess the likelihood of these intention errors.

The CREATE methodology involves two main stages: a modeling stage where CES is used to find situations that can lead to intention failures and therefore to erroneous actions; and a systems analysis input stage where the results of the cognitive modeling are integrated into the overall systems analysis.

The main steps in the modeling stage are:

- Decide what NPP situations to investigate with CES and how these situations map into the CES simulation world,
- 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,
- Analyze CES behavior to identify the minimum conditions which produce intention failures and the actions that follow from an intention failure.

Because CES is a simulation code, it requires detailed and complete input to run and outputs specific predictions about human intentions. This means that using CES in the modeling stage ensures explicit consideration and detailed analysis of the factors that contribute to human intention errors.

The main steps in the systems analysis input stage are:

- Modify the systems analysis event/fault trees to reflect the effects of intention errors identified in the modeling stage.

- Employ a quantification procedure to assess the likelihood of these intention errors,

- Combine the intention error estimates with execution error estimates.

Note that CES plays the same role in the CREATE methodology that simulation codes for physical plant processes play in reliability analyses of physical systems. In both cases we are dealing with complex, dynamic processes whose behavior is affected by too large a set of interacting factors to be tractable without a simulation. The modeling stage provides the backbone of the analysis in that it defines the critical elements to be aggregated and how they are to be aggregated. Frequency estimation techniques are then used to establish the probabilities to be aggregated.

1.3 Background for Model Development

This section briefly describes the background for the model development work carried out in Phase II including the goals to be satisfied, the behavioral science and NPP scopes to be addressed, and what activities are to be modeled. NUREG/CR-4532 contains a thorough discussion of these topics.

<u>Objectives of Model Development.</u>
The goal of the Phase II model development was to enhance the ability to predict human performance in NPPs, in particular, to enhance the ability:

- to predict the human contribution to risk in human reliability analysis (HRA) and probabilistic risk assessment (PRA);

- to identify situations prone to human error, particularly human related common mode errors and errors of commission;

- to understand the mechanisms that produce human error;

- based on increased knowledge about error mechanisms, to help develop error and risk reduction strategies;

- to predict the effects of changes in the NPP person-machine system (procedures, training, sensors, displays, operator aids) on human performance.

Intention Errors and Cognitive Processing.
Model development focused on one part of human behavior: human intention formation (deciding what to do) and erroneous intentions to act. Intentions to act are formed based on reasoning processes that determine how plant data are monitored, what situation assessments are formed, what explanations are built and what responses are judged appropriate to carry out under these perceived circumstances. The scientific disciplines that study these processes are called cognitive sciences. Models of these processes are often called "cognitive models."

What is cognition? The word cognitive describes one approach to understanding human behavior which assumes that description, explanation, and prediction of observable human actions depends on understanding the chain of information processing or mental events that mediate between observable events in the world and human responses.

> Definition of Cognition: The cognitive approach asserts that human performance varies because of differences in the knowledge that a person or team of people possess (both the form and the content), in the activation of that knowledge, and in the expression or use of knowledge.

How knowledge is activated and used is based fundamentally on an iterative cycle of data-driven activation of knowledge and knowledge-driven observation and action. An item in the world is noticed (e.g., an alarm) which triggers some knowledge (e.g., what the message means about changes in system state); this knowledge, in turn, leads to new observations or actions which trigger other knowledge, etc. Cognitive models differ in the particulars of how data activate knowledge and how activated knowledge leads to particular observations and actions in different contexts.

Scope.
The model development addresses the cognitive processes that affect successful and erroneous human intention formation in NPP emergencies. This area of human behavior includes what is sometimes called "rule-based" behavior and "knowledge-based" behavior up to the point of creative problem solving (Rasmussen, 1986).

Model development focused on one part of the NPP: operations during abnormal and emergency conditions, i.e., activities carried out by the emergency response system including the control room and branching out to the technical support center.

What Cognitive Activities Need to be Modeled?

An effective cognitive model must be able to capture the kinds of cognitive activities that occur in emergency operations in order to produce valid predictions that are relevant to NPPs. Let's call this target the basic competencies of the desired cognitive model. These competencies are kinds of behavior or information processing the model must exhibit that reflect aspects of the processing that people carry out to meet the demands of problem solving during control room emergencies.

To build a model to do this we must know -- What kinds of problem solving situations occur in NPP emergency operations? What must people know and how must people use that knowledge to solve these problems? How do people actually respond in these types of problem solving situations? The answers to these questions come from current empirical and analytical results on the cognitive demands and activities that arise in emergency operations (these are described in Chapter 4 of NUREG/CR-4532).

There are four primary characteristics of the NPP world that determine the kinds of problem solving situations that can arise in emergency operations.

1. NPPs are composed of a large number of highly interactive parts and processes (systems, functions, goals).

2. Emergency operations occur in a dynamic, event-driven world where incidents unfold in time, and events can happen at indeterminate times during an incident.

3. There is uncertainty -- a demanded position indication may not reflect actual position or sensors can fail -- and there is risk -- possible outcomes can have large costs.

4. There is a high degree of automation which means that multiple agents (machine controllers, machine decision makers, and multiple people) are involved in the response to emergency incidents.

The result is that actual NPP emergency incidents are difficult because multiple, interacting events (machine and human failures) can and do occur in the face of uncertain evidence and risky choices.

To solve problems in a world with these characteristics, operators must know about the many parts and processes and their interrelationships. They must be able to use this knowledge in a changing situation to determine the state of the plant (sustained monitoring) and how to respond (e.g., take into account side effects). This is complicated by uncertainties in the available evidence and the likelihood of multiple faults. Because of the high workload, operators must make decisions about timesharing/scheduling of activities. Because the world can be constantly changing, the ability to revise one's assessment of the situation, current goal, and current response strategy in response to new information is basic to problem-solving in this domain. This means the cognitive activities of the operator are best modeled as being "opportunistic" or interruptable by perceived changes in the state of the world. Emergency response can crystallize into a situation where the operators must make a choice among response strategies based on an uncertain situation assessment and risky possible outcomes (e.g., as occurred during the Ginna steam generator tube rupture incident).

In summary, the cognitive demands of NPP emergency operations produce the following processing requirements:

- process evidence to build a situation assessment given the possibility of multiple failures,

- sustained monitoring of evidence because it is a dynamic changing world,

- only a portion of the available evidence or possible explanations can be examined or pursued at any point -- attentional focus -- because of high workload and limited mental resources,

- there must be control and revision of attentional focus because it is a dynamic changing world,

- attentional focus is controlled through an interactive cycle of opportunistic, interruptable processing of new signals or events and knowledge driven choices about where to focus next,

- choice under uncertainty and risk.

A Cognitive Model.

The principal aim of a model is to efficiently capture relations among significant variables in order to describe, explain, and predict the behaviors of interest. To do this, models contain *concepts* and relations among concepts which specify what is really important in producing and controlling behavior in the situation of interest.[3] The concepts suggest what to look at

[3] As Eddington (1939, p. 55) remarked, "in physics everything depends on the insight with which the ideas are handled before they reach the mathematical stage."

and how to describe the situations that arise. CES is based on concepts about how intentions are formed and how they go astray that are derived from specific studies of human performance in NPP emergencies and general results in cognitive psychology.

Second, models are *representations* of some aspects of the situation of interest. They do not duplicate the modeled world; there is a relation between the modeling system and the modeled system. CES is a modeling environment designed as a parallel world to actual emergency operations. CES translates from a description of the evolution of an incident and recovery responses in terms of NPP engineering language to a description in terms of a cognitive problem-solving language in order to identify difficult or error prone problem-solving situations.

Third, models have some *machinery* to formalize the concepts and to generate specific and reproducible outputs given some inputs. Concepts about the processes involved in intention formation require formalization as symbolic processing or AI mechanisms. CES was developed based on the knowledge representation and processing mechanisms of the EAGOL AI software system developed by H. Pople and Seer Systems.

The next chapter and Volume 2 describe the concepts and formalization machinery in CES. Volume 2 is intended for those who want to know about how CES models successful and erroneous human intention formation.

Finally, models have multiple uses. This model was developed in order to better capture the human contribution to risk in probabilistic risk assessment studies. The methodology for using CES in PRA is described in Chapter 3 of the executive summary and in Volume 3. Volume 3 is intended for those who are interested in how the modeling capabilities of CES can be utilized in HRA and PRA.

2. Modeling Intention Formation with Cognitive Environment Simulation

The feasibility study done in Phase I found that all attempts to provide causal models of human performance in worlds where a broad range of cognitive activities occur result in *framework* models (e.g., Pew & Baron, 1983; Baron, 1984; Pew et al., 1986; Mancini et al., 1986). Framework models use one kind of modeling concept or technique to build a structure for the different kinds of cognitive activities that occur in the domain of interest and to capture how they interact. Narrower scope modeling concepts derived from heterogeneous sources provide depth at different points in the structure. This modeling strategy is used in many domains because there is a tradeoff between the desire for a formal model and the need to cover a broad scope of human behavior when modeling complex technological worlds (see sections 2.5 and 3.2 of NUREG/CR-4532).

The framework for the modeling system developed in this research program is based on a *model of the problem-solving environment* that is emergency operations. The emphasis is first on modeling the cognitive demands imposed by the problem-solving environment (the nature of the emergency incident, how it manifests itself through observable data to the operational staff, how it evolves over time?). Then, concepts from narrower scope psychological models (monitoring dynamic systems, e.g., Moray, 1986; choice under uncertainty and risk, etc.) can be brought to bear to represent the factors that affect human behavior in meeting these demands and to constrain the model of the problem-solving environment. The most fundamental psychological constraint relevant to the NPP world is that people have limited cognitive processing resources, and this cognitive model was designed to simulate a *limited resource problem solver in a dynamic, uncertain and complex situation.*

Because this modeling approach was chosen, the resulting modeling capability has been named a *Cognitive Environment Simulation or CES.*

CES is a *causal model* in the sense that it generates predictions about operator action by simulating the processes by which intentions are formed. This contrasts with correlational approaches that base predictions on descriptive regularities between situational variables (e.g., time available to respond) and performance (e.g., likelihood of making an error) without simulating the processes that produce the error. The ability to simulate the processes that lead to a particular intention makes it possible to predict likely behavior in complex and dynamic situations where operator intentions depend on a large number of interacting factors (e.g., what data he has available, number of issues competing for his attention, what he knows about the meaning of that data, the order that different kinds of explanations come to mind that could account for the data) that would otherwise be intractable. Furthermore, it enables identification of the form of the error (e.g., a fixation error) and the sources of the error (what aspects of the situation confronting the operator and/or of his knowledge or cognitive processing limitations, contributed to the error.)

CES is formally expressed as an AI based computer problem solving system that carries out cognitive processes that are critical to intention formation in complex dynamic worlds -- it monitors plant behavior, forms a situation assessment, generates one or more explanations for the plant state, forms expectations as to the future course of plant behavior (e.g., that automatic systems will come on or off), and generates intentions to act. In particular, CES is a specific instance of the EAGOL artificial intelligence problem solving system that is capable of reasoning in complex dynamic worlds (See footnote 2.) Among EAGOL's unique strengths are the ability to reason in multiple fault situations and to reason in situations that evolve over time (i.e., where evidence accrues over time, where evidence may disappear or become occluded by new events, where beliefs about the state of the world must be revised, etc.).

Degrading these capabilities or what we call the basic cognitive competencies of CES, leads to error vulnerable problem solving behavior. Poor performance -- errors -- emerges from a mismatch between demands (the incident) and the knowledge and processing resources. Varying CES knowledge and processing resources increases or decreases the program's vulnerability to getting offtrack or, once offtrack, staying offtrack. In this view, errors are the outcome of a processing sequence, and a model of error mechanisms depends on a model of processing mechanisms. Thus, the cognitive activities that underlie the formation of an intention to act are encompassed in CES and errors arise due to limitations of these cognitive processes. This is the imperfect rationality approach to modeling human performance and error (e.g., Rasmussen, Duncan & Leplat, 1987).

Modeling consists of matching CES resources to those present in some actual or hypothetical NPP situation. The specific processing mechanisms in CES are not intended to be "micro" models of human cognitive processing. It is the outcome of the computer's processing activities that are assumed to be the same -- what data are monitored, what knowledge is called to mind, what situation assessment is formed, what explanations are adopted, and what intentions to act are formed, *given* the incident (the demands of the problem-solving situation), the representation of the world (i.e., as reflected in the displays by which the operator interacts with the world), and the set of knowledge and processing limitations set up in CES.

The CES modeling environment provides powerful facilities for exploring how what a person knows, what data about the world are available to him, and his monitoring and problem-solving strategies can lead to successful or unsuccessful performance in different dynamic situations. Users of the model can express different particular NPP situations by selecting the demands (the incident or variant on the incident) and by adjusting the resources within the simulation to analyze and predict "what would happen if."

2.1 Cognitive Environment Simulation

CES is a dynamic simulation capability for human intention formation. As shown in Figure 2-1, CES takes as input a time series of those values that describe plant state which are available or are hypothesized to be available to be looked at by operational personnel. Any valid source of data about how the plant would behave in the incident of interest can be used to create the inputs to CES. This includes data on plant behavior in actual incidents or simulation results derived from training simulation models, engineering simulation models, or nuclear-thermohydraulic codes.

The dynamic stream of input data constitutes a *virtual display board* which the CES simulation monitors to track the behavior of the plant over time, to recognize undesirable situations, and to generate responses which it thinks will correct or cope with these situations (intentions to act). Its output is a series of these intentions to act which are then executed and therefore modify the course of the incident.

CES is a modeling environment for the *supervisory role* during emergency operations. This is because CES does not actually execute its intentions. Another mechanism is needed to actually carry out CES's instructions on the power plant. For example, a person, who has access to controls to a dynamic plant simulation, can execute CES instructions. Whether this person executes CES's instructions correctly or not depends on the nature of the incident which the CES user wishes to investigate.

CES watches the virtual display board of potentially observable plant behaviors and generates actions that it thinks will correct or cope with the perceived situation. To do this, inside of CES there are different kinds of processing which are carried out "in parallel" so that intermediate results established by one processing activity can be utilized by another and visa versa. This allows a solution to be approached iteratively from different levels of analysis.[4] There are three basic kinds of activities that go on inside of CES (Figure 2-2):

- Monitoring activities -- what parts of the plant are tracked when; are observed plant behaviors interpreted as normal-abnormal or expected-unexpected?

- Explanation building activities -- what explanations are considered, in what order, and adopted to account for unexpected findings?

- Response management activities -- selecting responses, either expected automatic system or manual operator actions, to correct or cope with observed abnormalities, monitoring to determine if the plans are carried out correctly, and adapting pre-planned responses to unusual circumstances.

An analyst can look inside CES to observe these activities as the incident it was stimulated with unfolds in time. The analyst can see what data the computer simulation gathered, what situation assessments were formed, what hypotheses were considered, pursued or abandoned, what plant behaviors were expected or unexpected. This can be done interactively, assuming that CES is being stimulated by a dynamic plant simulation and assuming that CES intentions are being executed on the simulated plant. Or an analyst can examine a record or description of the knowledge activated and processed by CES after it has been stimulated by an incident. In both cases CES's processing activities and resulting intentions to act are available to be analyzed (1) to identify erroneous intentions, (2) to look for the sources of erroneous intentions, (3) to discover what other actions follow from erroneous intentions (Figure 2-3).

The CES user can vary the demands placed on CES -- how difficult are the problems posed by the input incident. The CES user also varies the resources within CES for solving the problems by modifying what knowledge is available and how it is activated and utilized. The dimensions along which CES performance can vary are called *CES Performance Adjustment Factors* (or PAFs). There are a variety of these adjustment factors designed into CES that provide tools for a human analyst to set up or model the particular NPP situations which he or she wishes to investigate within the cognitive environment simulated in CES. For example, CES should be

---

[4] In some psychological models there are linear stages of information processing where an input signal is processed through a fixed sequence of stages. In CES, different processing occurs at the same time and intermediate results are shared. This leads to formalisation as an AI program.

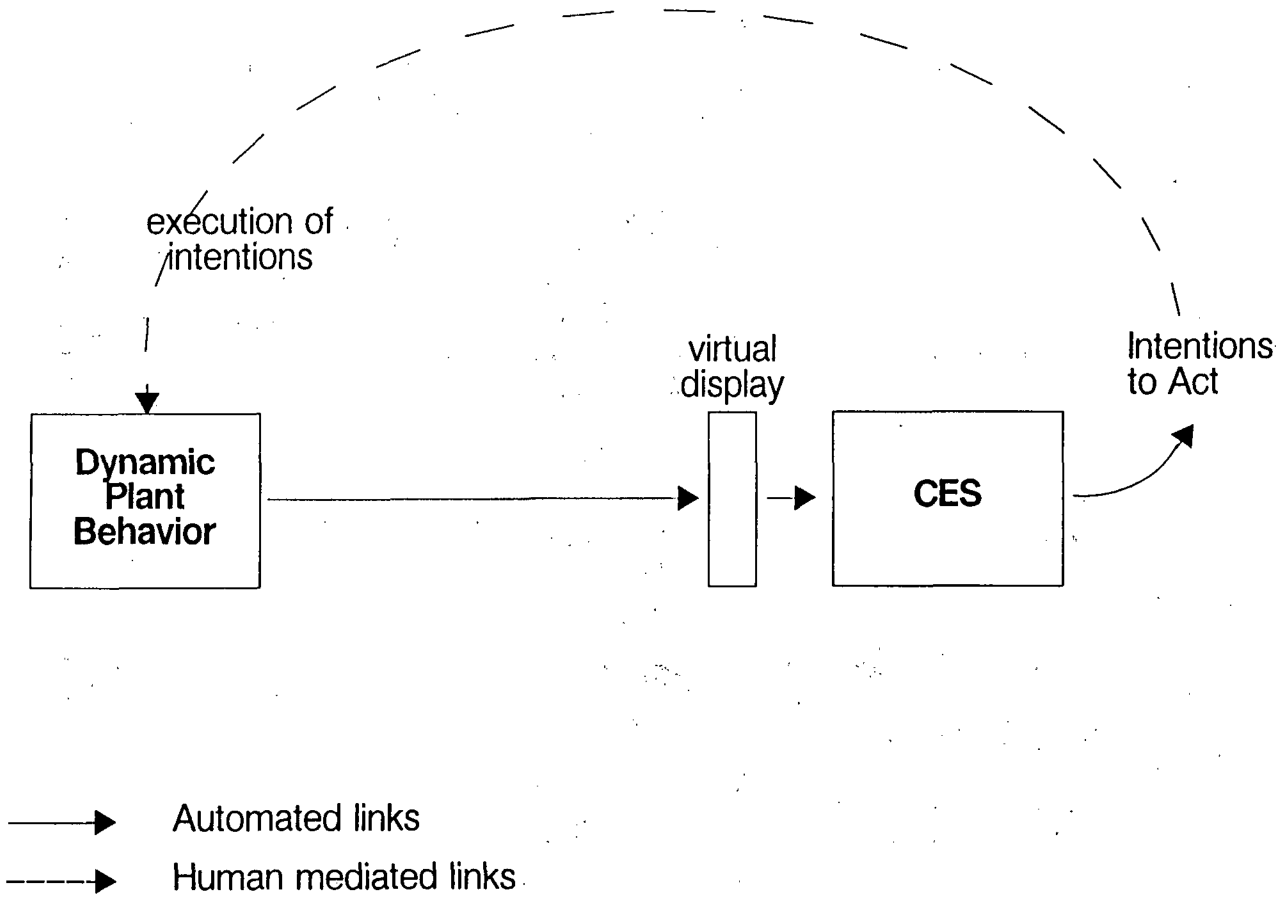


Figure 2-1: CES is a dynamic simulation capability for human intention formation. It takes as input a time series of those values that describe plant state which are available or are hypothesized to be available to be looked at by operational personnel. The CES simulation watches this virtual display board of potentially observable plant behaviors to track the behavior of the plant over time, to recognize undesirable situations, and to generate responses which it thinks will correct or cope with these situations (intentions to act). Its output is a series of these intentions to act which are then executed and therefore modify the course of the incident.

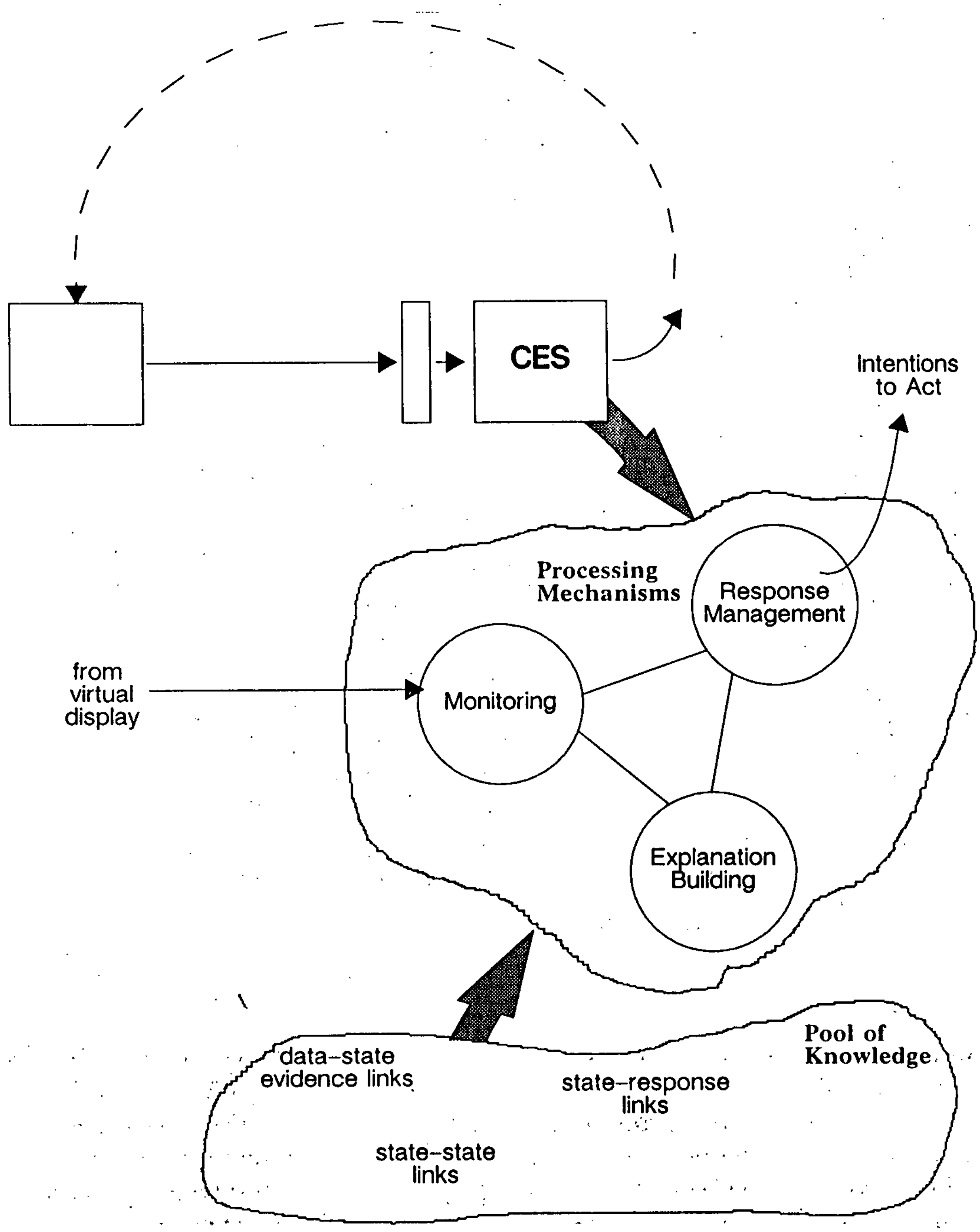


Figure 2-2: Inside of CES there are different kinds of processing which are carried out "in parallel" so that intermediate results established by one processing activity can be utilized by another and visa versa. This allows a solution to be approached iteratively from different levels of analysis. There are three basic kinds of activities that go on inside of CES: (a) monitoring activities -- what parts of the plant are tracked when and are observed plant behaviors interpreted as normal-abnormal or expected-unexpected? (b) explanation building activities -- what explanations are considered, in what order, and adopted to account for unexpected findings? (c) response management activities -- selecting responses, either expected automatic system or manual operator actions, to correct or cope with observed abnormalities, monitoring to determine if the plans are carried out correctly, and adapting pre-planned responses to unusual circumstances.

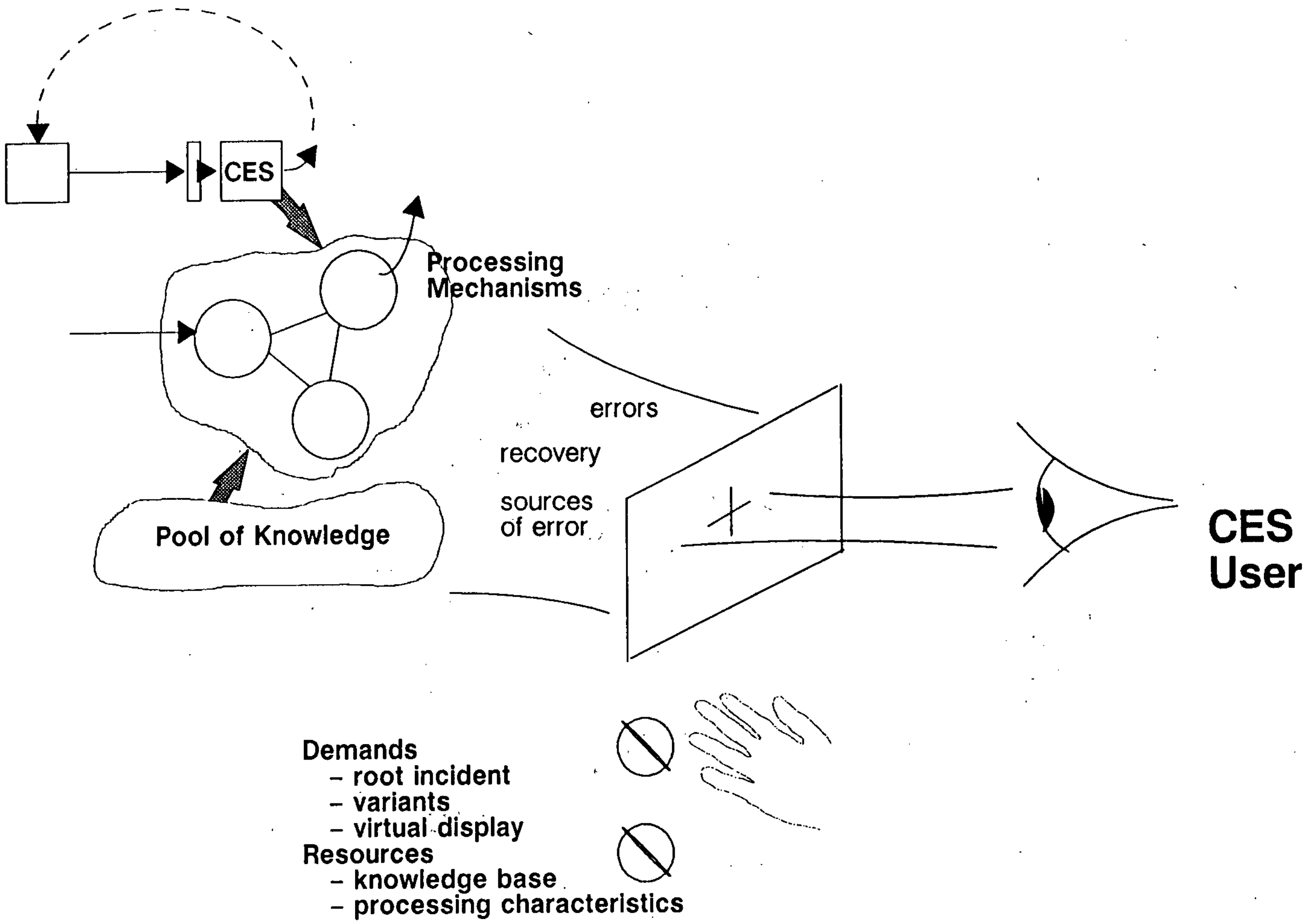


Figure 2-3: An analyst can look inside CES to observe these activities as the incident it was stimulated with unfolds in time. The analyst can see what data the computer simulation gathered, what situation assessments were formed, what hypotheses were considered, pursued or abandoned, what plant behaviors were expected or unexpected. CES's processing activities and resulting intentions to act are available to be analyzed (1) to identify erroneous intentions, (2) to look for the sources of erroneous intentions, (3) to discover what other actions follow from erroneous intentions.

The CES user varies the demands placed on CES -- how difficult are the problems posed by the input incident. The CES user also varies the resources within CES for solving the problems by modifying what knowledge is available and how it is activated and utilized. The dimensions along which CES performance can vary are called CES Performance Adjustment Factors (or PAFs).

capable of responding in a "function-based" and/or in an "event-based" fashion to faults, and CES should be capable of being fixation prone or not being fixation prone in explanation building. Modeling NPP situations within the CES simulation environment is, in effect, a translation from the engineering languages of NPP incidents to a problem solving language as represented by the knowledge and processing mechanisms set up in CES. CES is then run to find the conditions that lead to erroneous intentions and the action consequences of these erroneous intentions.

A human performance model must be built based on knowledge of what people actually do in the situations of interest. If one knew this completely, then the benefit of formal modeling is to eliminate subjectivity in the application of this knowledge to specific cases. But our knowledge of human performance in complex dynamic worlds such as NPP operations is incomplete (e.g., Hollnagel, Mancini & Woods, 1986). Given this state of affairs, formal models are needed (a) to objectively express the current state of knowledge, (b) to extrapolate from this to new situations, (c) to test whether the current state of knowledge is adequate through comparisons to new empirical cases, and (d) to revise and update the state of knowledge as appropriate (the model as a repository of current knowledge/best guesses/approximate models on operator behavior).

The Cognitive Environment Simulation allows one to formally represent the state of knowledge about what people do in emergency operations (or alternative views about what they do) and then to see the implications of that knowledge (or point of view) for human intention formation in new situations where there is no or sparse empirical data. Thus, a cognitive environment simulation allows one to generate analytical data on human performance that complement, but do not replace, empirical data on human performance.

This state of affairs is analogous to the situation with analytical computer codes which model reactor behavior. In both cases, an ongoing cycle of model evolution and change is needed as our state of knowledge changes. The Cognitive Environment Simulation, as repository of the best current knowledge, then, becomes the best source for interpolating or extrapolating what human behaviors are likely in cases where there is no or limited experience -- including evaluating changes to the human-machine system and hypothetical situations that arise in postulated incidents for which there is no or insufficient empirical data (rare incidents). Reactor thermodynamic models are essential tools for design and risk assessment of the physical NPP. The Cognitive Environment Simulation provides, for the first time, an analytical model of human intention formation in NPP emergency operations which will be an essential tool to assess human performance for the evaluation of human-machine systems in the NPP and for assessment of the human contribution to risk.

2.2 CES Development Process

The process by which CES was created is illustrated in Figure 2-4.

Concepts and relations about how intentions are formed and how they go astray derived from empirical results and knowledge about the structure and function of NPPs were used to formulate a set of basic *cognitive competencies* that CES should exhibit. As mentioned earlier, the basic competencies are imposed by the need to simulate a limited resource problem solver in a dynamic, uncertain and complex situation.

If CES was to function as a modeling environment, the cognitive competencies also needed to include the dimensions along which CES performance should be variable -- CES Performance Adjustment Factors (PAFs). CES should be capable of competent performance given some set of Performance Adjustment Factor settings, and should be capable of incompetent performance given other Performance Adjustment Factor settings. Furthermore, the performance breakdowns which CES exhibited under different PAFs must be related to what is known about how human problem-solving can break down in dynamic situations.

The concepts about intention formation were derived from general results in cognitive psychology and from empirical studies of human performance in NPP emergencies (cf., Chapter 4 of NUREG/CR-4532). Empirical results used included both studies of operators solving simulated faults (Woods et al., 1982; Woods & Roth, 1982; and unpublished cases) and retrospective investigations of operator decision making in actual incidents (e.g., the four incidents analyzed in Pew et al., 1981; the Ginna and Oconne incidents analyzed in Woods, 1982 and Brown & Wyrick, 1982; the Davis-Besse incident reported in NUREG-1154; the San Onofre incident reported in NUREG-1190; the Rancho Seco incident reported in NUREG-1195).

In the CES development process, the different types of knowledge that a person might possess about the NPP were also taken into account. CES had to be capable of representing these different kinds of knowledge and different ways of organizing knowledge about the NPP. The formalism for organizing knowledge about NPPs that informed CES development is based on Gallagher et al. (1982), Woods & Hollnagel (1987), and Woods (in press).

A formalization process followed where AI mechanisms embodied in the

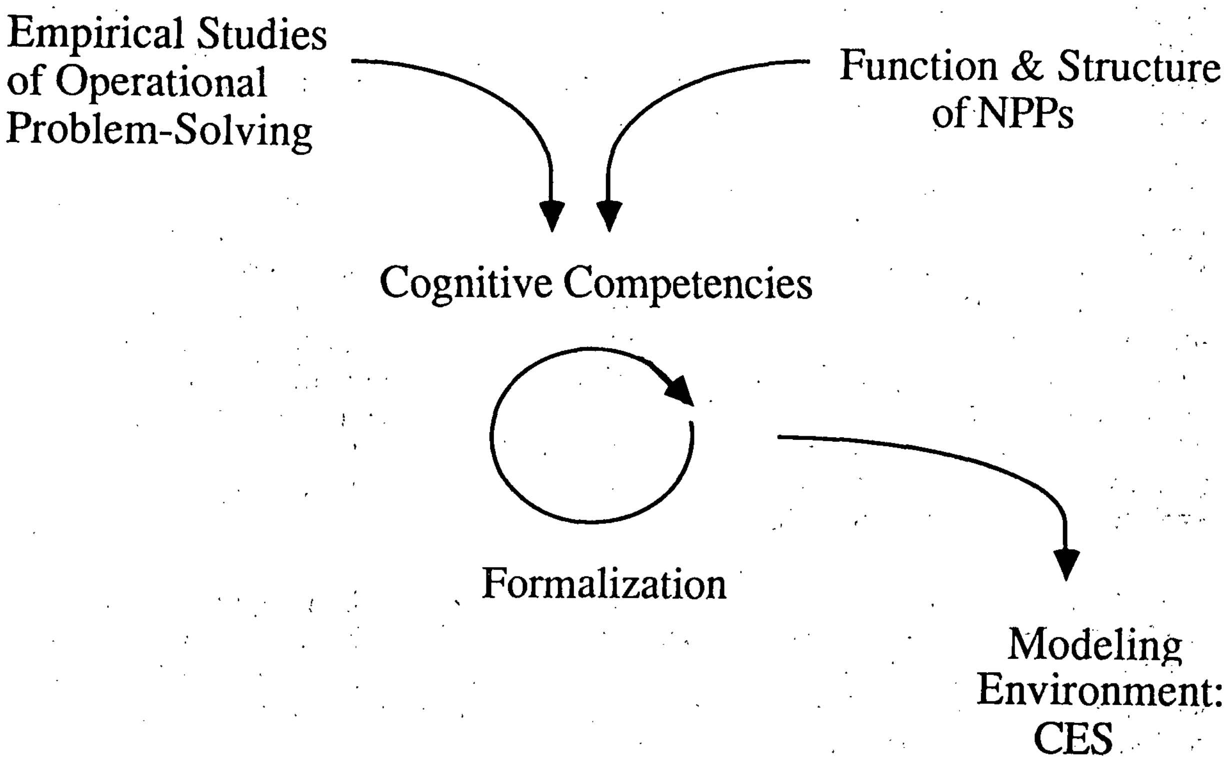


Figure 2-4: The CES development process. Concepts and relations about how intentions are formed and how they go astray were used to formulate a set of basic cognitive competencies that CES should exhibit. A formalization process followed where AI mechanisms were set up that could exhibit those competencies. Several iterations of formalization, leading to more refined statement of the basic competencies and then further formalization were carried out to develop CES to its current state.

EAGOL artificial intelligence problem solving system were set up that could exhibit those competencies. The basic software mechanisms had to be capable of competent performance and capable of being degraded to exhibit the kinds of performance breakdowns that humans exhibit in high cognitive demand situations. Several iterations of formalization, leading to more refined statement of the basic competencies and then further formalization were carried out to develop CES to its current state.

Practical limitations meant that some aspects of the ideal scope of NPP tasks and cognitive activities were not considered in the initial development. CES currently exhibits only a part of all of the target cognitive competencies. However, the goal was to capture enough of existing empirical and theoretical knowledge about operator cognitive activities in emergency situations for CES to begin to be a useful tool to explore what would people do if situation x arose and to identify situations prone to intention failures. The process of using CES will then provide useful information on human performance and reliability at the same time that it undergoes further evolution, extensions and refinement.

The mechanisms for interacting with CES (setting up plant input, modifying model performance adjustment factors) are currently very limited, as are the mechanisms for watching and recording CES behavior when it is stimulated by dynamic sequence of plant data. These mechanisms are easily expandable to improve ease of use and analyst productivity.

2.3 Overview of the CES Architecture

As an instance of the EAGOL AI computer system, CES contains two major types of information. First, it contains a *knowledge base* that represents the operator's (or the team of operators') knowledge about the power plant, including the inter-relationships between physical structures, how processes work or function, goals for safe plant operation, what evidence signals abnormalities, and actions to correct abnormalities.

Second, it contains *processing mechanisms* (or inference engine) that represents how operators process external information (displays, procedures) and how knowledge is called to mind under the conditions present in NPP emergencies (e.g., time pressure). This part of the model determines what knowledge is accessed when and what cognitive activities (monitoring, explanation building, response management) are scheduled when during an evolving incident.

The knowledge representation formalism from EAGOL (i.e., how knowledge about the NPP is expressed) provides a powerful and flexible mechanism for representing virtually any relation among NPP concepts. Concepts at any level of abstraction, whether observable or not, can be represented (e.g, a plant parameter reading; an intermediate disturbance category such as a "mass imbalance"; a fault category such as primary system break to containment; or a response such as "turn off the emergency cooling system"). Within the knowledge representation formalism, the full variety of relations among concepts that NPP operators would be expected to know such as plant data-state evidence links, state-state links, and state-response links can be expressed. This includes encoding of symptom-response "shortcuts" that form the basis for what has sometimes been termed operator "rule-based" behavior, as well as encoding of more abstract and functional relations that form the basis for more elaborated reasoning or what has sometimes been termed "knowledge-based" behavior (Rasmussen, 1986).

Included in the knowledge representation is a description of what data about plant state are directly available to the model to "see," reflecting what plant information would be directly available to the operator to observe. This description constitutes a *virtual display board*, that the model monitors to acquire plant data information. The CES knowledge base includes a list of plant parameters or states that it can directly access (e.g., from a data file or as output from a simulation program). Depending on the plant being modeled these plant parameters can be direct sensor readings, or more integrated information about plant state such as the output of computerized displays or decision aids). Associated with each element on the "virtual display board" are parameters that reflect characteristics of how that information is presented in the plant being modeled (i.e., characteristics of the *representation* provided to the operator of that NPP).

The basic psychological concept behind CES is that people have limited resources in a potentially high workload environment. This means that CES, as a model of operational personnel, cannot access and utilize all possibly relevant pieces of knowledge (i.e., not all potentially relevant knowledge in the knowledge base can be *activated*) on any one model *processing cycle*, i.e., time step). Similarly, CES cannot examine all of the plant data available at any one processing cycle. Therefore, CES must be able to control what data are examined when and what knowledge (and how much knowledge) is activated in a given cycle. This is one of the basic cognitive competencies specified for CES.

Controlling what knowledge and how much knowledge is activated at a given point in an unfolding incident depends on:

- A cycle or interaction between *knowledge-driven processing* (such as looking for information to find an explanation for an unexpected finding) and *data-driven processing* (where salient data interrupt ongoing processing and shift the focus).

- *Resource/workload interactions* where carrying out one type of processing precludes the possibility of doing other processing if there is competition for limited resources. Thus, there can be a need to choose which processing activity should be carried out next, e.g., acquire more data? or pursue possible explanations? or generate/track responses to detected abnormalities?

- A limited problem solver should focus first on *"interesting" findings*. There are several layers of criteria that define which findings are "interesting" or "important" that affect control of CES processing. For example, if an observation indicates an abnormality, then there is a need to pursue how to correct or cope with it; if an observation is unexpected, then there is a need to pursue what could account for it?

The formalization task then was to use the symbolic processing or AI mechanisms in EAGOL to control a limited focus of attention in these ways, e.g., what data are examined when, what possible explanation is pursued first.

The basic processing mechanism from the EAGOL system used in CES to achieve this behavior is to spawn an "analyst" when some criterion is met, who then performs some information processing work, accessing knowledge available in the knowledge base as it needs it. There are three basic kinds of "analysts" each with their own area of responsibility and with different criteria that trigger their processing activities. These are:

- *Behavior Analysts* responsible for monitoring and analyzing plant behavior to decide if observed plant behaviors are expected or unexpected.

- *Situation Analysts* responsible for analyzing the perceived situations and for postulating and pursuing possible explanations for unexpected findings.

- *Response Plan Analysts* responsible for selecting and adapting plans to correct or cope with perceived abnormal conditions.

These analysts are active processes that draw conclusions and "post" their results for other analysts to use as needed. Multiple instances of each basic type of "analyst" are generated or "spawned" as needed. A fundamental characteristic of this problem-solving architecture is that each analyst has a very narrow field of view and responsibility, and that complete problem solving involves communication and coordination among the multiple analysts.

Each analyst does not represent a different person, rather the cooperative set of analysts are intended to model a single problem-solving system -- be it an individual operator or a team of operators. The multiple analysts are intended to model the multiple types of processing (e.g., monitoring, explanation building, response planning) and lines of reasoning (e.g., multiple alternative explanations pursued) that occur in parallel and are interwoven during problem-solving

2.4 Modeling Human Intention Formation with CES

CES, as a modeling environment, is designed as a parallel world to actual emergency operations. The parallel is established by capturing in the simulation world the problem solving resources available in some actual or hypothetical NPP situation (operators, training, procedures, control board, etc.). If the parallel is well established, then the behavior of the simulation (the monitoring, explanation building and response management behavior of CES) in some incident corresponds to expected human behavior in the actual world, under the same circumstances.

2.4.1 Changing CES Resources: Performance Adjustment Factors

CES is a deterministic model. Given the same dynamic incident scenario, the same virtual display board characteristics, the same knowledge about the NPP, and the same processing resources, CES will generate the same series of intentions to act. There are large degrees of variability in human behavior; even when performance is good, people take different trajectories to reach the same outcome. CES is capable of large degrees of variation in its behavior as well, and it is capable of taking different problem solving trajectories to the same outcome.

Variability in CES behavior arises from several sources. First, variability in CES behavior arises due to variability in details in how the incident in question unfolds. This is one reason why dynamic plant behavior is needed as input to CES. Second, CES behavior varies as a function of variations in its knowledge and processing resources. The assumption is that human variability arises from differences in relatively enduring knowledge (e.g., knowledge of how x works) and processing characteristics (e.g., a fixation prone personality), longer term changes in knowledge and skill (e.g., skill acquisition from training or experience), or from more moment-to-moment variations in processing resources (e.g., a narrow field of attention due to stress or fatigue).

There are a set of factors designed into CES which allow one to vary CES knowledge and processing resources. These Performance Adjustment Factors (PAFs) provide the tools for a human analyst to establish parallels between the cognitive environment simulated in CES and NPP situations which he or she wishes to investigate. The analyst uses PAFs to represent the resources available (or thought to be available) in a particular NPP situation within the CES modeling environment. The CES user then stimulates CES with data on plant behavior in different incidents, checks how CES solved those problems (intention failures, omission and commission errors that follow from intention failures, error recovery), re-adjusts PAFs to explore variants, and re-runs CES to identify the conditions under which intention errors occur, the consequences of intention errors, and the sources of intention errors.

There are a variety of adjustment factors designed into CES which provide the tools for a human analyst to establish parallels between the cognitive environment simulated in CES and NPP situations which he or she wishes to investigate. The analyst can use these Performance Adjustment Factors (PAFs) to model a particular NPP situation within the CES modeling environment, stimulate CES with data on plant behavior in different incidents, check how CES solved those problems (intention failures, omission and commission errors that follow from intention failures, error recovery), re-adjust PAFs to explore variants, and re-run CES to identify the conditions under which intention errors occur, the consequences of intention errors, and the sources of intention errors.

Traditional performance shaping factors (e.g., experience level, stress, organizational climate) are examples of variables that are thought to affect human behavior. CES Performance Adjustment Factors (PAFs) are variables that affect CES behavior. To simulate a NPP situation in CES, the factors operative in that situation which are thought to affect human behavior are mapped into CES PAFs in a *two step inference process.* First, one must specify what is the impact of the factor of interest (or a change in that factor) on cognitive activities. This can be derived from theoretical concepts (e.g., the effect of team structure on problem-solving processes), empirical data, or analysis. In any case, it is the effects on the *processes* involved in activating and utilizing knowledge which must be specified. Second, the specified effects on cognitive processing are translated into adjustments in PAFs.

For some kinds of performance shaping factors this two stage inference process is a straightforward tractable analytical task. For example, with respect to the effect of procedures on performance, the specific guidance on corrective responses encoded into the procedures (e.g., specification of corrective responses to take) would be extracted and entered into the CES knowledge base. With respect to issues of display quality, the relative salience of different plant data on a control board or in a computer display system would be determined by the CES user analytically and used in the set up of the virtual display board.

Other kinds of factors can be specified based upon straightforward empirical investigation. For example with respect to effects of training or experience one can use simple "quick and dirty" techniques or more sophisticated techniques to find out what particular operators actually do know about how some plant process works (e.g., natural circulation), about the basis for some response strategy, or about what possible hypotheses are brought to mind by some plant behavior(s).

Finally, some factors require a specification of *how they are assumed to affect problem-solving processes* in order to be mapped into CES PAFs. For example, how does stress affect problem-solving (e.g., high stress might narrow the field of attention) or how do different organizational structures affect problem-solving? The answer to this question specifies what PAF settings should be used to investigate the consequences of this factor on intention formation errors in different incidents and over various other PAF settings.

Note that the answer to this question requires taking a theoretical position on how the factor in question impacts on the processes involved in problem solving. The theoretical relation asserted can be derived from behavioral science research (e.g., the impact of team structure on problem solving) or analyst judgment. This is an example of how CES is a framework model that utilizes more specific models in some areas.

2.4.2 Cognitive Processing and Erroneous Intentions

An analyst uses CES PAFs to change the knowledge and processing characteristics within CES or the virtual display of data to CES. This allows the CES user to explore under what conditions intention failures occur and to see the consequences of intention failures for further actions on the plant in different incidents or variations on a root incident. Errors -- failures to form the appropriate intentions for the actual situation -- depend on how CES activates and uses knowledge, given the demands of the incident under investigation.

Finding intention failures with CES is based on the concept that the difficulty of any given problem-solving situation is a function of

1. *The problem-solving demands.*

    - Processing requirements imposed by the characteristics of the incident (e.g., a multiple fault incident where one masks the evidence for another is inherently more difficult to isolate that a single fault incident with a clear signature).

    - The *representation* or window on the world by which the problem-solver views and interacts with the incident (e.g, the displays available on the control board; integrated information available on computer-based displays).

2. The available *problem-solving resources.*

    - The base of knowledge about the NPP that is available to use in problem-solving. This includes knowledge about the structure and function of the NPP and knowledge about NPP disturbances/faults, and how to correct these.

    - The processing mechanisms and their characteristics (e.g., size of the field of attention, how fixation prone, degree of communication among different processing mechanisms).

Errors emerge when there is a mismatch between demands and resources. For example, a narrow field of attention (low resources) cannot lead to intention failures if the incident in question produces no situations where a wide field of view is needed for timely detection of important plant behaviors (low demands).

Intention formation errors are the end result of a processing sequence which starts and develops due to failures to call to mind relevant knowledge -- a plant behavior is missed (which could happen due to several factors, such as, because of low observability of the data or because the focus of attention is elsewhere), the knowledge it would have evoked is not brought to mind and does not lead to an accurate situation assessment (e.g., plant behavior x is interpreted as expected instead of unexpected), the erroneous situation assessment affects what explanations are pursued or not pursued and what responses are evoked or not evoked. Varying CES processing resources through PAFs increases or decreases the program's vulnerability to getting offtrack or, once offtrack, staying offtrack.

3. Cognitive Reliability Assessment Technique: CREATE

CES is a simulation of cognitive processes that 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, person-machine interfaces). By simulating the cognitive processes that determine situation assessment and intention formation, it provides the capability to establish analytically how people are likely to respond, given that these situations arise.

The process for using the CES computer simulation tool as part of PRA studies to enhance human reliability assessment is called Cognitive Reliability Assessment Technique or CREATE. In the CREATE methodology, the capabilities of CES are used to find situations where intention failures occur and to find the risk consequences of the erroneous actions that follow from the intention failure (errors of omission, errors of commission, common mode errors). This chapter provides an overview of how CES is used to find intention failures and the basic steps in the CREATE methodology. Volume 3 provides a more detailed account.

In the CREATE methodology, CES is used in a way that closely parallels how simulation codes for physical plant processes are currently used in PRA studies. This is because CES serves the same role in anticipating human behavior as other simulation codes serve in predicting the behavior of thermodynamic processes. In both cases we are dealing with complex, dynamic processes whose behavior is affected by too large a set of interacting factors to be tractable without a simulation.

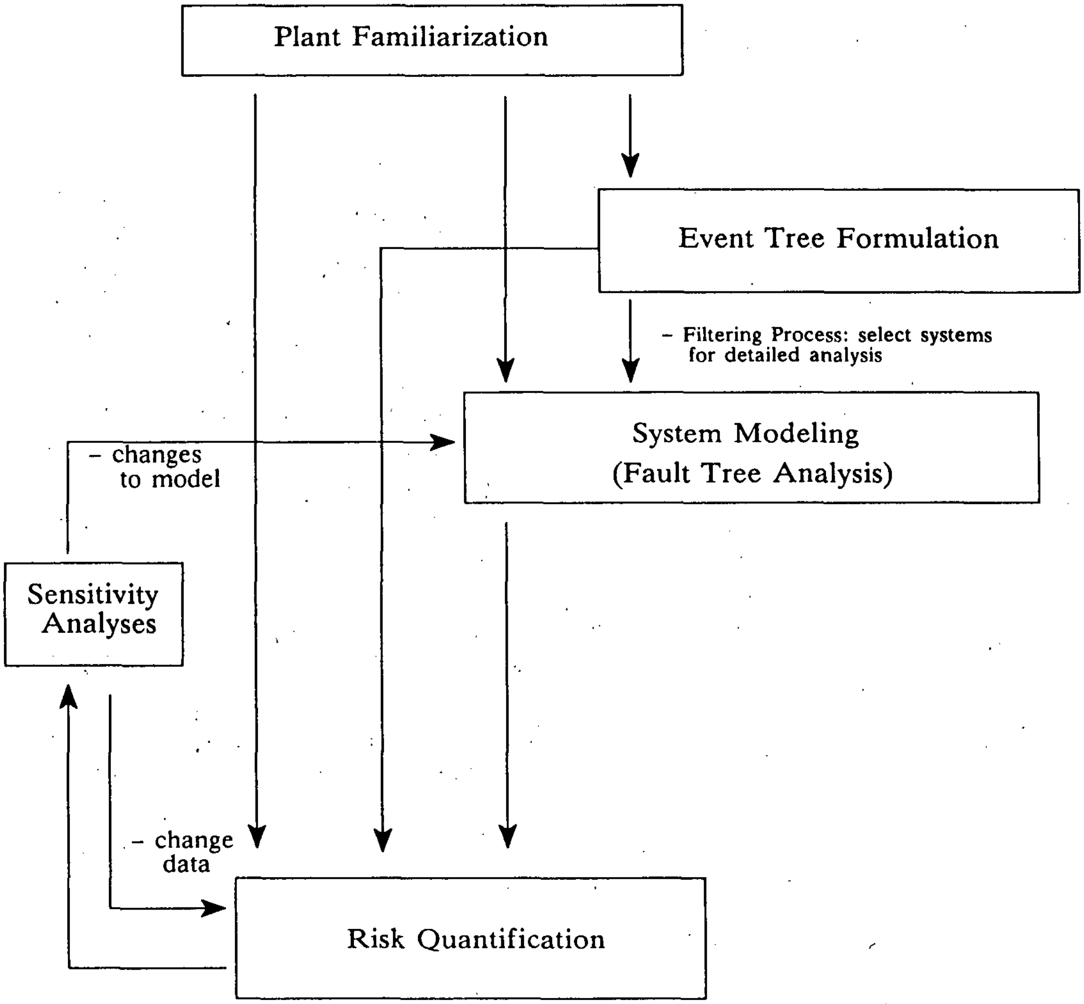
The CREATE methodology basically involves two major stages: a modeling stage where CES is used to find situations that can lead to intention failures and therefore to erroneous actions; and a systems analysis input stage where the results of the cognitive modeling are integrated into the overall systems analysis.

The major premise of the CREATE methodology is that predicting operator intentions and actions under accident conditions requires the same level of detailed modeling and sensitivity analysis that is performed to predict the behavior of physical processes in the plant (see NUREG-1150). Human reliability analysis has the same status as equipment component reliability analysis, with the two analyses proceeding in parallel and each drawing on the insights of the other.

Human reliability primarily enters into the systems analysis stage of a probabilistic risk assessment study. In this stage the dominant accident sequences by which plant components may fail and lead to core damage are identified. Figure 3-1 provides a block diagram of the systems analysis process.[5] Figure 3-2 contains a block diagram of the key steps in the CREATE methodology. Comparison of Figures 3-1 and 3-2 reveals the parallel between the current PRA approach to hardware reliability analysis, and the proposed process for analyzing human reliability. In both cases, analyses involve a modeling stage and a quantification stage. Both analyses begin with a plant familiarization procedure, proceed to an event tree formalization of failure sequences that can result in accident situations, and then go on to a detailed modeling step. These three activities taken as a whole constitute a modeling stage. The results of the modeling stage provide input to a risk quantification stage where estimates of risk are quantified probabilistically. While the stages of the analyses appear linear, there is in fact a good deal of interaction and iteration among the stages. In the CREATE methodology, hardware component reliability and human cognitive reliability are treated as separate, equal status analyses that merge to generate overall system reliability assessment. The two analyses are intended to proceed in parallel, but to closely interact and draw insight from each other.

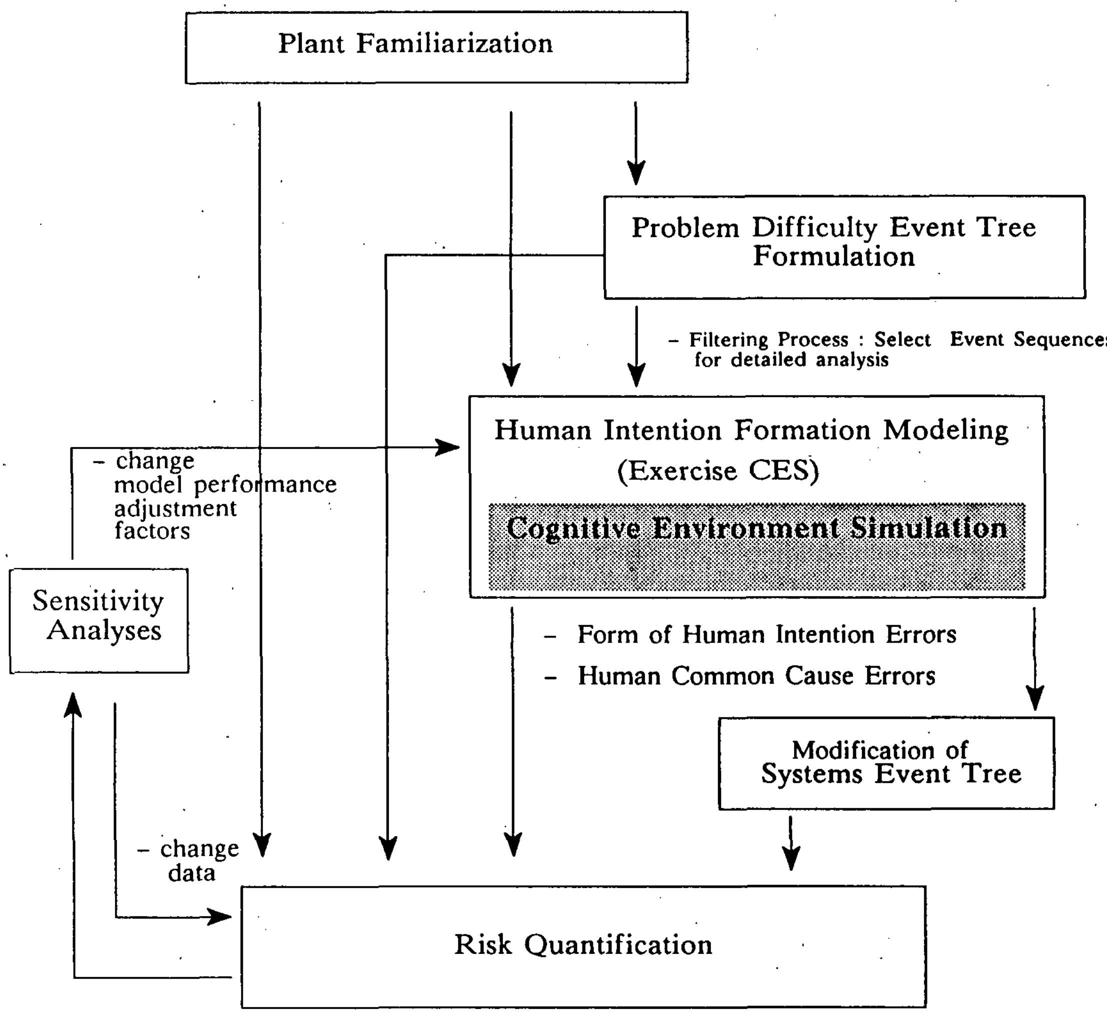
In both systems analysis and CREATE the modeling stage provides the backbone of the analysis, in that it defines the critical elements to be aggregated and how they are to be aggregated (e.g., independent/dependent, alternative paths that require adding probabilities, or conjunctive relations that require multiplying probabilities). In the systems analysis stage fault-trees are used to identify the alternative paths (minimum set of component failures or cut sets) that will result in a system failure. This defines which component failure frequencies need to be estimated and how they are to be aggregated to compute overall system failure probability. Similarly, CES is used to identify alternative sets of conditions (or, more precisely, sets of minimum conjunction of conditions) that will result in intention failures with risk significant consequences. This defines the conditions for which frequency estimates will need to be obtained, and how the estimates will need to be aggregated to produce an overall probability estimate of a given intention failure arising. The result of the CES modeling stage will also dictate changes to the systems analysis event/fault trees since it will identify cases of intention errors with multiple action consequences (e.g., errors of commission, common mode failures) that were previously unanticipated. This

[5] Figure 3-1 is adapted from a figure that appears in NUREG 1150, which is the most recent NRC sponsored probabilistic risk assessment study.



Process for Hardware Oriented Systems Reliability Analysis

Figure 3-1: Block diagram of the systems analysis stage of a probabilistic risk assessment study (adapted from a figure in in NUREG 1150). During the systems analysis stage the dominant accident sequences by which plant components may fail and lead to core damage are identified. Note that the systems analysis process involves a systems modeling phase (event tree and fault tree analyses) followed by a quantification phase. In the CREATE methodology a parallel analysis process involving a modeling and quantification stage is employed in assessing human reliability (see Figure 3-2).



Cognitive Reliability Assessment Technique (CREATE)

Figure 3-2: Block diagram of the key steps in the Cognitive Reliability Assessment Technique or CREATE. Comparison of Figures 3-1 and 3-2 reveals the parallel between the current PRA approach to hardware reliability analysis, and the process for analyzing human reliability outlined in CREATE. In both cases, analyses involve a modeling stage and a quantification stage. Both analyses begin with a plant familiarization procedure, proceed to an event tree formalization of failure sequences that can result in accident situations, and then go on to a detailed modeling step. These three activities taken as a whole constitute a modeling stage. The results of the modeling stage provides input to a risk quantification stage.

will allow exploration of new branches on the systems analysis event tree and/or adjustment of the probability estimates assigned to existing nodes.

3.1 Stages in the CREATE Methodology

3.1.1 Overview

The CREATE methodology requires that an HRA analyst participate as part of the PRA team in the Systems Analysis stage. The HRA analyst on the PRA team should have expertise in the behavioral sciences. This is necessary to ensure that he or she can bring to bear the empirical and theoretical knowledge of human problem-solving in complex dynamic worlds such as NPPs, in interpreting the behavioral implications of the plant-specific data that he/she collects, and mapping characteristics of the plant operating environment into the CES simulation environment.

One of the earliest steps in a PRA systems analysis is a plant familiarization stage. This involves a review and analysis of plant-specific conditions and reliability data. In the CREATE methodology, the HRA analyst would participate actively during this stage to collect plant-specific data necessary to set up and run the CES model and to estimate likelihood values for input to the quantification stage.

During the familiarization process the HRA analyst will engage in a variety of activities designed to gather plant-specific data on operating conditions and philosophy that affect human performance during emergency conditions (e.g., characteristics of control room displays and controls; contents of operator training and procedures). The Team-Enhanced Evaluation Method (TEEM), developed under a previous NRC program, provides a detailed methodology for the kinds of mock-ups, walk-throughs, detailed interviews, site visits, and iterative refinement of task analyses that are involved in an enhanced human reliability analysis process (O'Brien, Luckas, & Spettell, 1986).

While the HRA analyst's activities during plant familiarization (e.g., detailed interviews, walk-throughs, review of procedures, task analyses) are largely similar to the activities traditionally prescribed for comprehensive human reliability analyses, the data collection process will be strongly structured and guided by the detailed data requirements for setting up and running CES, and data requirements for quantifying the frequency of different processing demand and resource conditions that serve as input to the quantification stage.

The results of plant familiarization forms the basis of the modeling stage. The analyst uses information gathered during the plant familiarization and plant task analyses to set up the general parallel between this plant and CES, to set up any particular parallels he or she wishes to investigate with CES (different incidents, variations in the virtual display to CES, variations in knowledge and processing resources), and then to actually investigate what are likely human behaviors in these situations. The basic steps in this process are:

- Decide what NPP situations to investigate with CES and how these situations map into the CES simulation world,
- Set up CES to be able to run NPP situations, i.e., enable CES to accept and process plant data during the incidents and variants of interest and represent within CES the results of plant specific analyses or hypotheses about knowledge and processing resources which analysts wish to investigate,
- Run CES over a plausible range of demand and resource settings, given the analysis of this plant,
- Analyze CES behavior to identify the minimum conditions which produce intention failures and the actions that follow from an intention failure.

It is important to keep in mind that these are not a linear series of steps; there is a great deal of interaction between these steps in the modeling stage.

The modeling results are used to generate inputs to the systems reliability analysis. During this stage,

- the errors identified are used to modify and enrich the event trees in the plant systems reliability analysis,
- a quantification procedure is used to assess the likelihood of these intention errors,
- intention error estimates are combined with execution error estimates.

3.1.2 Modeling Stage

In order to use CES to find intention failures one must decide what NPP situations are promising or important to investigate, including different incidents and the factors that are likely to affect human behavior such as traditional performance shaping factors. An important aspect of the modeling stage in CREATE is the identification of high cognitive demand situations to investigate using CES.

One starting point for identifying accident scenarios to be explored with CES, are the accident or event sequences selected by PRA analysts for in depth analysis. These accident sequences are derived from the event tree analyses, and are defined by an accident initiating event (e.g., Loss of Offsite Power, Loss of Coolant Accident), and a series of postulated subsequent events (e.g., additional failures of systems necessary to respond to the initiator or its consequences).

These accident sequences, which we will refer to as *root incidents*, are likely to be underspecified with respect to features of the situation that impact on information processing and problem-solving that are critical for assessing how an operator is likely to perceive the situation and form intentions to act. For example, from a problem-solving point of view, whether control room displays are functioning correctly or not (e.g., due to sensor failure; loss of electric power) can have profound impacts on performance, although PRA event trees are typically not specified at that level of detail. Consequently, a critical element of the CREATE methodology is to fill in and define variants on the root incident that represent good candidates for in-depth analysis with CES (i.e., high problem-solving demand situations that are likely to produce human errors of intention.)

In the CREATE methodology the HRA analyst develops a *problem difficulty event tree* that defines variations on the root incident that increase the difficulty of the problem-to-be-solved because they degrade the ability to perform necessary tasks or because they impose additional tasks. The kinds of incident to investigate are ones where there is some variation or difficulty that goes beyond the standard method for handling the situation, or *complicating factor*, such as:

- human execution errors,
- additional machine failures (e.g., valves that stick open, systems that fail to work as demanded),
- missing information (e.g., sensor failures),

- multiple major faults (tube rupture with an unisolatable steam release from the faulted steam generator),

- situations which remove or obscure the usual 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),

- complex situations where different parts of the situation suggest responses which conflict with each other (e.g., the Ginna incident),

- situations that require actions that depart from the usual (e.g., total loss of feedwater).

The incidents selected for analysis with CES will be those variants on the root incident that are identified to be high problem-solving demand situations.

In addition to selecting incidents to stimulate CES with, the HRA analyst needs to map characteristics of the plant operating environment into the CES simulation world. This involves:

- finding what operational personnel know or might know about the incidents in question (e.g., from training or procedures) and entering that information into the CES knowledge base,

- setting up the CES virtual display board to model characteristics of control board displays, alarms, and computer-based displays available to be looked at by operational personnel,

- setting up CES processing mechanisms to represent settings or hypotheses about how people may activate and utilize knowledge under different conditions which analysts wish to investigate.

Carrying out these steps currently requires intimate knowledge of CES and its computer implementation.

Multiple runs of each incident variant are performed varying the representation of the state of the plant and problem solving resources assumed to be available to the operator based on analyses of the plant in question. The goal of the parametric investigation is to identify situations where mismatches arise between the problem-solving demands imposed by the situation and the problem-solving resources available to the operator.

The record of how CES attempted to solve each incident is then analyzed to identify the minimum conditions which produce intention failures and to identify the actions or non-actions that follow from each intention failure:

- What actions will not be attempted – errors of omission.

- What actions the intention failure will lead to – *commission errors*, "improvised" action sequences (that are not the response sequence specified in available procedures) and *common mode failures*. Information about these errors may result in changes to event trees on the plant side of risk analysis (e.g., adding a branch to the original event tree) or signal a need to switch to or create a different event tree.

- The likelihood of detecting and recovering from errors – human intention failures or execution errors or failures of plant equipment to respond as demanded. Of particular value, it will provide information on the likely time duration before an erroneous action is detected and recovery begins.

CES is deterministic in the same sense that any one calculation of a simulation of a process produces deterministic results. The probabilistic element is the likelihood of the circumstances arising which led to the intention error committed by CES. This is analogous to the process of computing the health consequences of a postulated accident given uncertainty in what the prevailing weather patterns will be at the time of the accident. Meteorological models are available to assess the radiation effects for any given weather pattern. The major element of uncertainty relates to the probability of different weather patterns at the time of the accident.

3.1.3 Inputs to System Reliability Analysis

In the modeling stage, CES is used to find situations that can lead to intention failures and the resulting erroneous actions. A second stage in the CREATE methodology is to incorporate these results into the risk analysis. The sources and results of intention errors will be merged with the results of the hardware reliability analyses that are going on in parallel to generate overall system reliability estimates. The results of the CREATE modeling stage will affect overall systems reliability analysis in two ways: (1) it will lead to modifications to the systems event tree analysis and (2) it will provide quantitative estimates of human reliability as input to overall system reliability frequency estimates.

One of the major contributions of the CES modeling stage is that it will reveal events in the systems event tree that need to be expanded or modified. In the process of in depth exploration of potential intention errors, forms of intention errors that have one or more action consequences (e.g., errors of commission, common mode failures) that were not anticipated in generating the initial version of the systems event tree will be identified. This will allow addition of a new branch to the systems analysis event tree and/or adjustment of the probability estimates assigned to existing events. Finding intention failures can indicate the need for modification of reliability estimates for several systems assumed to be independent but found to be linked via an operator error of intention formation (i.e., identification of human-related common mode errors).

The CES modeling stage can also point to situations where operators are likely to initiate actions, not explicitly prescribed, that lead to recovery of failed systems and reduce overall risk. The identification of plausible operator initiated recovery actions provide an objective basis for expanding the systems event tree to include the potential recovery action.

It should be stressed that while, for expository purposes, input to systems analysis is placed late in the CREATE process, in practice there will be a great deal of interaction and mutual feedback between the human reliability analysis and the hardware reliability analysis throughout. The PRA process in general involves a considerable degree of iteration and interaction among stages, and the systems event trees are modified and expanded as new insights are gained from related analyses. The CREATE methodology has the HRA analyst actively participating in the systems event tree formulation process, contributing insights gained from the CREATE process as they are revealed.

The procedure for estimating intention failure likelihoods (cognitive reliability estimates) from CES assumes that the major element of uncertainty in predicting operator behavior rests on assessing the probability that the situation will arise which produces intention errors, when that situation is simulated with CES. As a first approximation, it can be assumed that *any* limited resource intelligent agent would exhibit the intention error produced by CES with a probability approaching one, when placed in the same situation (i.e, given the cognitive demands imposed by the incident, the representation characteristics, and the problem-solving resource limitations).

Estimating intention failure likelihoods primarily involves assessing the probability that the situation will arise which was found to result in an intention failure (as identified by running CES). There are two components to the situation: one is how likely is it for these problem solving demands to arise and the second is how likely is it that a particular set of problem solving resources will be available or in effect. For example, exercising CES may indicate that given a particular accident scenario, and a particular set of PAF settings (e.g., narrow field of attention) an intention error will occur. The questions for estimation are (1) the likelihood that the characteristics of the accident scenario that contributed to the intention error will arise, and (2) the likelihood that operator cognitive resources will match the PAF settings (e.g., that operators' field of attention would be narrow).

The first component, the probability of the cognitive demanding situation arising, is estimated by existing systems reliability techniques. The second component, the probability of the particular set of problem solving resources being in effect, requires behavioral science expertise or data (e.g., studies to determine what actual operators know about how system x functions).

In the absence of empirical data on these questions, estimation will require expert judgment. The estimation tasks should be easier to perform and provide more accurate results than typical expert estimation tasks because CES results in a shift in the nature of the events whose frequencies are to be estimated. CES enables classes of events that occur with relative frequency to be substituted for rare events for which frequency information is unobtainable. This is because CES reformulates the question of the probability of an intention error into the question of the probability of the conditions that lead to an intention error arising. The former question requires judgment of the probability of gross behaviors (i.e., an intention failure) under hypothetical conditions that go beyond people's base of experience, for which, even in principle, accurate frequency information is beyond reach. The latter question requires frequency estimates to be made for classes of events (i.e., conditions that increase cognitive demands) that occur with higher frequency and for which a base of experience has been developed. Because there is a relevant experiential base to draw upon, it is possible to gather the empirical data or to generate more accurate estimates from expert judgment.

In most cases the HRA analyst will have collected sufficient relevant data during plant familiarization to be in a position to estimate frequencies without needing to convene a panel of experts (e.g., what proportion of operators have high experience and what proportion are inexperienced?).

Note that some aspects of a situation such as the representation of plant state or the guidance in procedures or computer advisory systems available to operational personnel are generally deterministic. They can be determined for a particular plant by HRA analysts on the Systems Analysis team during the plant familiarization and task analysis phases of the PRA.

Overall human reliability estimates can be obtained by combining the probability of human intention errors with the probability of human execution errors. Current theories of human intention errors and human execution errors postulate very different underlying generating mechanisms. As a result, at this stage of development it can be assumed that intention errors and execution errors are independent. As such, the joint probability of a failure to take a required action is obtained by combining disjunctive likelihood estimates.

It is important to note that a large degree of interaction between the human analysis and the plant systems analysis is required in CREATE. The two analyses proceed in parallel each posing questions to the other (what is likelihood the human will do x? what is the likelihood the difficult variation on root incident y will arise?) and each providing input to the other analysis (human commission and common mode errors; what incident variations increase problem difficulty).

4. Conclusions and Recommendations

4.1 Benefits of the CREATE Methodology

The leverage provided by the CREATE methodology revolves around the unique capabilities provided by the CES cognitive model. These capabilities overcome many of the limitations in human reliability analysis that have been long recognized by the PRA community. Among the benefits of this approach are better modeling of the sources of human-related risk, more accurate quantification of the level of human-related risk, and deeper insights into conditions that produce significant human errors.

CES is the first cognitive process simulation tool that allows exploration of plausible human responses in different emergency situations. By simulating the cognitive processes that determine situation assessment and intention formation, it provides the capability to establish analytically what actions an operator is likely to take under different accident conditions. This means one can investigate the ability of humans to *recover* from equipment failures, execution errors or intention failures to stop or mitigate their consequences. Similarly, one can investigate *errors of commission* due to misperception of plant state or other cognitive error.

The ability of CES to predict errors of commission is particularly important since misapprehension of plant state by the operator can result in multiple actions which can have broad systemic effects. Intention failures are a major source of *human related common mode failures* -- multiple failures that are attributable to a common element (namely, the erroneous intention). For example, cases where the situation is misperceived, and the operator deliberately decides it is appropriate to turn off multiple, otherwise redundant and diverse systems as occurred at Three Mile Island and Chernobyl. The PRA community generally recognizes the importance of identifying common mode failure events because they can have large and widespread effects on risk. CES represents the first cognitive process model able to predict the wide spread consequences that can follow from an intention failure.

The CES computer simulation provides risk analysis with a tool for deriving likely human responses in different situations. In the past most of the tools available for human reliability analysis have taken the form of guidelines and checklists. While these tools provide useful guidance, there is a large subjective component in identifying contributors to human error and the likely form of human response under accident conditions. As a result, there has been wide variability among PRA studies in assessment of human impact on risk (Worledge, Chu and Wall, 1984). CES is a simulation code that requires detailed and complete input to run and outputs specific predictions about human intentions. Because CES requires detailed input to run, it ensures explicit consideration and detailed analysis of the factors that contribute to intention errors.

There are other benefits that derive from the modeling capabilities of CES. One can investigate the sources of cognitive processing breakdowns and intention failures. Because CES encompasses the factors that effect the available problem solving resources such as the specific form and content of displays, training, and procedures, it provides an analytic tool for investigating the effects of changes in NPP person-machine systems including new instrumentation, computer-based displays, operator decision aids, procedure changes, training, multi-person or multi-facility (e.g., technical support center) problem solving styles. This means that risk reduction strategies can be evaluated. Similarly, proposed changes/enhancements to NPP person-machine systems can be analytically checked before they have been implemented. The cognitive model can be used to filter which changes are sufficiently likely to improve performance that prototype construction and empirical tests are justified. As such it should provide a cost-effective complement to difficult and expensive high-fidelity empirical evaluations.

CES/CREATE has the promise to become an essential tool in the assessment and regulation of systems that affect the human element in NPP safety.

4.2 Recommendations

The next steps which are needed to take advantage of the capabilities of the CES cognitive model and the CREATE methodology are:

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

The usefulness of the CES cognitive model and the CREATE methodology depends on the ability of CES to behave like people do, for the same situation and with the same external and internal resources. In other words, the key question to be answered is the validity of CES as a modeling tool for human intention formation. An initial empirical evaluation and validation study is planned for Phase III of the research project.

Analogous to the situation with analytical computer codes which model reactor behavior, there needs to be an ongoing cycle of model evolution and change as our state of knowledge changes. The Cognitive Environment Simulation is the repository of the current state of knowledge on operator cognitive activities and is the best source for interpolating or extrapolating what human behaviors are likely in cases where there is no or limited experience – including evaluating changes to the human-machine system and hypothetical situations that arise in postulated incidents for which there is no or insufficient empirical data. To fulfill this function CES needs to evolve as new empirical data are gathered and as our understanding of human error evolves.

The current implementation of CES does not exhibit all of the target cognitive competencies specified for CES, and it addresses only a small portion of the ideal scope of NPP tasks. The full range of cognitive competencies needs to be incorporated into CES and the NPP scope covered by CES needs to be broadened.

The mechanisms for interacting with CES (setting up plant input, modifying Performance Adjustment Factors) are currently very limited, as are the mechanisms for watching and recording CES behavior when it is stimulated by dynamic sequence of plant data. As a result, at this stage of development CES can be effectively used only by people who have behavioral science expertise, particularly in cognitive processes and human error, and intimate knowledge of the AI computer structures used to implement CES (i.e., the EAGOL software system). Mechanisms for interacting with CES can be expanded and enhanced to improve productivity and accessibility.

4.3 Conclusion

As a result of the model development work in Phase II of this research project, there exists, for the first time, a simulation model of the cognitive processes that affect operator intention formation in NPP emergencies. Reactor thermodynamic models are essential tools for design and risk assessment of the physical NPP. Similarly, the CES cognitive model will be an essential tool to assess human performance for the evaluation of human-machine systems in the NPP and, via the CREATE methodology, for assessment of the human contribution to risk.

Enough knowledge about operator cognitive activities in emergency situations and enough knowledge about parts of the NPP have been incorporated for

CES to begin to be a useful tool to explore what would people do in NPP situations of interest and to identify situations prone to intention failures. The process of using CES will then provide useful information on human performance and reliability at the same time that CES undergoes further evolution, extensions and refinement.

Utilizing the capabilities of the CES cognitive model and the CREATE methodology in PRA studies requires changes in the relationship between human reliability and systems reliability analysis. The two analyses need to proceed in parallel each drawing on the insights of the other.

5. References

Baron, S. A control theoretic approach to modeling human supervisory control of dynamic systems. In W. B. Rouse (Ed.), *Advances in Man-Machine Research, Volume 1.* JAI Press, 1984.

Brown, W. & Wyrick, R. (eds.). *Analysis of Steam Generator Tube Rupture Events at Oconee and Ginna.* Institute of Nuclear Power Operations, 1982. (82-030).

Gallagher, J. *Disturbance Analysis and Surveillance System Scoping and Feasibility Study.* Palo Alto, CA: Electric Power Research Institute, 1982. (NP-2240).

Hollnagel, E., Mancini, G. & Woods, D. D. (Eds.). *Intelligent Decision Support in Process Environments.* New York: Springer-Verlag, 1986.

Joksimovich, V. A Review of Plant Specific PRAs. *Risk Analysis*, 1984, *4*(4), 255-266.

Levine, S. & Rasmussen, N. C. Nuclear Plant PRA: How far has it come? *Risk Analysis*, 1984, *4*(4), 247-254.

Mancini, G. Modelling Human and Machines. In Hollnagel, E., Mancini, G. & Woods, D. D. (Eds.), *Intelligent Decision Support.* Germany: Springer-Verlag, 1986.

Moray, N. Monitoring behavior and supervisory control. In K. Boff, L. Kaufman & J. Beattie (Eds.), *Handbook of Human Perception and Performance.* NY: Wiley, 1986.

O'Brien, J. N., Luckas, W. J., Jr., & Spettell, C. M. *Team-Enhanced Evaluation Method (TEEM) Procedures: An Enhanced Human Reliability Analysis Process* (Informal Report BNL-38585). Brookhaven National Laboratory, December 1986.

Pew, R. W. et al. *Cockpit Automation Technology* (Tech. Rep. 6133). BBN Laboratories Incorporated, 1986.

Pew, R. W. & Baron, S. Perspectives on Human Performance Modelling. In G. Johannsen & J. G. Rijnsdorp (Eds.), *Analysis, Design and Evaluation of Man-Machine Systems.* London: Pergamon Press, 1983.

Pew, R. W., Miller, D. C. & Feehrer, C. E. *Evaluation of Proposed Control Room Improvements Through Analysis of Critical Operator Decisions.* Palo Alto, CA: Electric Power Research Institute, 1981. (NP-1982).

Pople, H. Jr. Evolution of an Expert System: from Internist to Caduceus.

In I. De Lotto and M. Stefanelli (Ed.), *Artificial Intelligence in Medicine.* Elsevier Science Publishers B. V. (North-Holland), 1985.

Pople, H. E., Jr. Heuristic Methods for Imposing Structure on Ill-Structured Problems: The Structuring of Medical Diagnostics. In P. Szolovits (Ed.), *Artificial Intelligence in Medicine.* Boulder, CO: Westview Press, 1982.

Rasmussen, J. *Information Processing and Human-Machine Interaction: An Approach to Cognitive Engineering.* New York: North-Holland, 1986.

Rasmussen, J., Duncan, K. & Leplat, J. *New Technology and Human Error.* Chichester: John Wiley & Sons, 1987.

Reason, J. & Mycielska, K. *Absent Minded? The Psychology of Mental Lapses and Everyday Errors.* Englewood Cliffs, NJ: Prentice-Hall, 1982.

Swain, A. D. & Guttmann, H. E. *Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Applications.* Springfield, VA: National Technical Information Service, 1983. (NUREG/CR-1278).

Trager, E. A., Jr. *Case Study Report on Loss of Safety System Function Events* (Report AEOD/C504). Office for Analysis and Evaluation of Operational Data, U. S. Nuclear Regulatory Commission, December 1985.

U. S. Nuclear Regulatory Commission. *Loss of Main and Auxillary Feedwater at the Davis-Besse Plant on June 9, 1985.* Springfield, VA: National Technical Information Service, 1985. (NUREG-1154).

U. S. Nuclear Regulatory Commission. *Loss of Power and WAter Hammer at San Onofre, Unit 1 on November 21, 1985.* Springfield, VA: National Technical Information Service, 1985. (NUREG-1190).

U. S. Nuclear Regulatory Commission. *Loss of Integrated Control System Power and Overcooling Transient at Rancho Seco on December 26, 1985.* Springfield, VA: National Technical Information Service, 1985. (NUREG-1195).

U. S. Nuclear Regulatory Commission. *Reactor Risk Reference Document* (Tech. Rep. NUREG-1150). Office of Nuclear Regulatory Research, U. S. Nuclear Regulatory Commission, February 1987. (Draft for Comment).

Woods, D. D. Operator decision making behavior during the steam generator tube rupture at the Ginna nuclear power station. In W. Brown & R. Wyrick (Eds.), *Analysis of Steam Generator Tube Rupture Events at Oconee and Ginna.* Institute of Nuclear Power

Operations, 1982. (Also Westinghouse Research and Development Center Report: 82-1C57-CONRM-R2).

Woods, D. D. Coping with Complexity: The Psychology of Human Behavior in Complex Systems. In L. P. Goodstein, H. B. Andersen & S. E. Olsen (Eds.), *Mental Models, Tasks and Errors: A Collection of Essays to Celebrate Jens Rasmussen's 60th Birthday.* London: Taylor & Francis, in press.

Woods, D. D. & Hollnagel, E. Mapping Cognitive Demands in Complex Problem Solving Worlds. *International Journal of Man-Machine Studies*, 1987, *26*, 257-275. (Special Issue on Knowledge Acquisition for Knowledge Based Systems).

Woods, D. D. & Roth, E. *The Role of Cognitive Modeling in Nuclear Power Plant Personnel Activities: A Feasibility Study.* Washington D. C.: U. S. Nuclear Regulatory Commission, 1986. (NUREG-CR-4532).

Woods, D. D. & Roth, E. Operator Performance in Simulated Process Control Emergencies. Unpublished study, 1982.

Woods, D. D., Wise, J. A. & Hanes, L. F. *Evaluation of Safety Parameter Display Concepts.* Palo Alto, CA: Electric Power Research Institute, 1982. (NP-2239).

Worledge, D. H., Chu, B. B., & Wall, I. B. Nuclear Plant Systems Analysis Research at EPRI. *Risk Analysis*, 1884, *4*(4), 299-311.

NRC FORM 335 (2-84) NRCM 1102, 3201, 3202

U.S. NUCLEAR REGULATORY COMMISSION

BIBLIOGRAPHIC DATA SHEET

SEE INSTRUCTIONS ON THE REVERSE.

1. REPORT NUMBER (Assigned by TIDC, add Vol. No., if any)
NUREG-CR-4862
Volume 1

2. TITLE AND SUBTITLE
Cognitive Environment Simulation: An Artificial Intelligence System for Human Performance Assessment. Volume 1: Summary and Overview

3. LEAVE BLANK

4. DATE REPORT COMPLETED

MONTH	YEAR
September	1987

5. AUTHOR(S)
D. D. Woods, E. M. Roth, H. Pople, Jr.

6. DATE REPORT ISSUED

MONTH	YEAR
November	1987

7. PERFORMING ORGANIZATION NAME AND MAILING ADDRESS (Include Zip Code)
Westinghouse Research and Development Center
1310 Beulah Road
Pittsburgh, PA 15235

8. PROJECT/TASK/WORK UNIT NUMBER

9. FIN OR GRANT NUMBER
FIN D1167

10. SPONSORING ORGANIZATION NAME AND MAILING ADDRESS (Include Zip Code)
Division of Reactor and Plant Systems
Office of Nuclear Regulatory Research
U.S. Nuclear Regulatory Commission
Washington, DC 20555

11a. TYPE OF REPORT
Technical

b. PERIOD COVERED (Inclusive dates)
May 1986 through June 1987

12. SUPPLEMENTARY NOTES

13. ABSTRACT (200 words or less)
This report documents the results of Phase II of a three-phase research program to develop and validate improved methods to model the cognitive behavior of nuclear power plant (NPP) personnel. In Phase II a dynamic simulation capability for modeling how people form intentions to act in NPP emergency situations was developed based on techniques from artificial intelligence. This modeling tool, Cognitive Environment Simulation or CES, simulates the cognitive processes that determine situation assessment and intention formation. It can be used to investigate analytically what situations and factors lead to intention failures, what actions follow from intention failures (e.g., errors of omission, errors of commission, common mode errors), the ability to recover from errors or additional machine failures, and the effects of changes in the NPP person-machine system.

The Cognitive Reliability Assessment Technique (or CREATE) was also developed in Phase II to specify how CES can be used to enhance the measurement of the human contribution to risk in probabilistic risk assessment (PRA) studies.

The results are reported in three self-contained volumes that describe the research from different perspectives. Volume 1 provides an overview of both CES and CREATE. Volume 2 gives a detailed description of the structure and content of the CES modeling environment and is intended for those who want to know how CES models successful and erroneous intention formation. Volume 3 describes the CREATE methodology for using CES to provide enhanced human reliability estimates. Volume 3 is intended for those who are interested in how the modeling capabilities of CES can be utilized in human reliability assessment and PRA.

14. DOCUMENT ANALYSIS – a. KEYWORDS/DESCRIPTORS
Artificial Intelligence
Human Reliability
Probabilistic Risk Assessment
Cognitive Model
Nuclear Power Plant
Human Error
Problem Solving
Human Factors
Cognitive

b. IDENTIFIERS/OPEN-ENDED TERMS
CES
Cognitive Environment Simulation
Cognitive Reliability Analysis Technique
CREATE

15. AVAILABILITY STATEMENT
Unlimited

16. SECURITY CLASSIFICATION
(This page) Unclassified
(This report) Unclassified

17. NUMBER OF PAGES

18. PRICE

☆U.S. GOVERNMENT PRINTING OFFICE:1987-202-292:60287

UNITED STATES
NUCLEAR REGULATORY COMMISSION
WASHINGTON, D.C. 20555

OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE, $300

SPECIAL FOURTH-CLASS RATE
POSTAGE & FEES PAID
USNRC

PERMIT No. G-67

FOR HUMAN PERFORMANCE ASSESSMENT