NASA Risk Management Program: Roles and Responsibilities Handbook

Appendix: Tools and Techniques

Revised Draft for Review

May 1989

National Aeronautics and Space
Administration

July 1993

PREFACE

This document delineates representative major tools and techniques expected to be employed in the NASA Risk Management Program. These delineations expand upon the descriptions of risk modeling and qualitative and quantitative risk-based decision-making in NASA's hazardous activities that were presented in the NASA Safety Risk Management Program Plan, Volumes I and II, April and June 1987. This document is an appendix to the handbook on Risk Management Roles and Responsibilities (NHB ______________) in which the context for the application in NASA of the tools and techniques described here is defined.

The tools and techniques discussed are generally advanced in nature and the reader's familiarity with their basics is assumed. References are provided as assistance. References are also provided for additional details and further developments of the tools and techniques presented.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

LIST OF FIGURES

TABLE OF CONTENTS (Continued)

LIST OF FIGURES (Continued)

LIST OF TABLES

TABLE OF CONTENTS (Continued)

LIST OF TABLES (Continued)

CHAPTER 1
INTRODUCTION

1.1 PURPOSE

This handbook provides a compendium of advanced analytical tools and techniques for tailored applications in support of risk management decision-making in NASA. The handbook is part of a set of NASA Management Instructions, handbooks, and supporting documents being developed at NASA Headquarters to aid NASA program; facility; and Safety, Reliability, Maintainability, and Quality Assurance managers in risk-based safety decision-making.

1.2 SCOPE

Methods for both qualitative and quantitative analyses are included in this handbook. The analyses encompass all phases of risk-based decision-making, from hazards identification to risks disposition*. However, the analysis methods for hazards identification are standard in the system safety process and are described in detail in NASA's system safety documentation, and so are only briefly touched on here to indicate their role in initiating the complete risk management process. The tools and techniques described in this handbook are presented tutorially to a considerable extent, but since they are generally advanced methods, the user of this handbook is assumed to be familiar with their basic elements. Useful textbooks and handbooks are referenced with which these basic elements can be reviewed. In addition, books and articles are referenced which provide sources for added details or further extensions of the methods summarized here. These methods focus mainly on qualitative and quantitative logic modeling for system or other application failure occurrence analysis, the classical and Bayesian development of data for the quantitative modeling, and procedures for qualitative or quantitative decision-making on the disposition (acceptance, tolerance, or mitigation) of the risks implied by the system failure occurrence analysis and associated analyses of the consequences of failures.

* See the Glossary for definitions of terms employed in this handbook.

1.3 DEFINITIONS

See the Glossary at the end of the handbook.

1.4 RELATED DOCUMENTS

a. NMI 8070.4, *Risk Management Policy for Manned Flight Programs*, January 1988.

b. (Draft) NMI 8070.X, *Risk Management Policy for Unmanned Flight Programs*, June 1988 (Under Review).

c. (Draft) NMI 8070.Y, *Risk Management Policy for R&T Facilities*, June 1988 (Under Review).

d. NHB __________, *NASA Risk Management Program: Roles and Responsibilities* (Main Volume) (Under Review).

e. *NASA Safety Risk Management Program Plan, V. I (The Program)*, June 1987, and Vol. II (*Rationale*), April 1987.

f. (Draft) *National Space Transportation System Risk Management Program Plan*, August 1987.

g. NHB 1700.1, Vol. 3, *System Safety*, 1968.

h. NSTS 22254, *Methodology for Conduct of NSTS Hazard Analysis*, May 1987.

i. (Draft) SSP 30309, *Safety Analysis and Safety Risk Assessment Requirements and Processes Document for the Space Station Program,* September 1988.

j. National Research Council, *Post-Challenger Evaluation of Space Shuttle Risk Assessment and Management*, January 1988.

k. U.S. Department of Defense, *Military Standard, System Safety Program Requirements*, MIL-STD-882B, March 1984.

l. Nuclear Regulatory Commission, *Probabilistic Risk Assessment Procedures Guide*, NUREG/CR-2300, 1983.

m. (Draft) NHB 1700.1, Vol. 7, *NASA System Safety Handbook,* 1988.

CHAPTER 2
SYSTEM SAFETY ANALYSIS AND THE RISK ASSESSMENT PROCESS

2.1 GENERAL DESCRIPTION

This handbook describes the qualitative and quantitative tools and techniques of risk management which can be applied in assessing and making decisions about the risks associated with NASA's hazardous activities. The background and delineation of the NASA Risk Management Program, and general descriptions of its functions, applicable methods, and responsibilities, were provided in the Safety Risk Management Program Plan, Volumes I and II (Related Documents 1.4.e). The present handbook provides greater detail on the methods, especially those of fault tree modeling, and of qualitative and quantitative risk-based decision-making.

Risk management consists of six main steps:

- identification of significant hazards, including potential faults and failures
- definition of possible accident sequences or scenarios due to those hazards
- qualitative assessment (e.g., categorization) or quantitative estimation of the frequency or probability of each accident sequence
- qualitative assessment (e.g, categorization) or quantitative estimation of the severity of the possible consequences and losses associated with each accident sequence
- combination of the frequency and consequence or loss assessments into qualitative assessments or quantitative estimates of risks, evaluation of their significance and deciding on their acceptability or need for mitigation, and evaluation of any needed mitigations
- tracking the efficacy of the acceptability or mitigation decisions.

The first and last of these steps are touched on only briefly in this manual. They are basic system safety engineering functions and are treated comprehensively in NASA's and others' system safety handbooks (e.g., Related Documents 1.4 g., j., and l.). The emphasis in the present manual is on the structured qualitative and quantitative risk management decision support methods that assist in the remaining four steps listed above.

As indicated in Table 2-1, there are numerous NASA activities that can pose potential hazards for NASA personnel and equipment as well as for the public at large. The most reliable source of information concerning accidents in such activities is direct or related experience. However, in many cases, particularly with low probability, high consequence accident events, little or no data are available. In this case, risk assessment techniques that include modeling and analysis must be used.

Uncertainty can arise in both qualitative and quantitative risk assessments. Uncertainty arises due to possible:

- incompleteness in modeling (all accident scenarios possibly not identified and all causes of these scenarios possibly not identified)

- need to make other than ideal modeling assumptions

- inadequacy of reliability data

- imperfect understanding of accident phenomenology

- difficulties in the assessments of consequences and losses.

An important part of the risk assessment process is the treatment of these uncertainties.

The risk assessment process, especially a probabilistic one, leads to insights not attainable from a deterministic, only consequence severity-oriented, analysis. These insights are emphasized in this manual.

MANNED LAUNCHES	
Flight Safety:	Hazards to Crew, Passengers, System, Mission, Payload
Range Safety:	Hazards to Public and NASA People and Property Debris and Explosive Fragments Radioactive Fragments and Releases Blast Toxic Emmissions Sonic Boom
UNMANNED LAUNCHES	
Flight Safety:	Hazards to System, Mission, Payload
Range Safety:	Hazards to Public and NASA People and Property Debris and Explosive Fragments Radioactive Fragments and Releases Blast Toxic Emmissions Sonic Boom
AERONAUTICS (Aircraft, Rockets, Balloons)	
Flight Safety:	Hazards to Crew, Passengers, System, Mission, Payload
Range Safety:	Hazards to Public and NASA People and Property Crash Impact Fire and Explosion Sonic Boom
TRANSPORTATION (All Modes)	
Accidents:	Hazards to Public and NASA People and Property Hazardous Materials Large Objects
STORAGE FACILITIES	
Accidents:	Hazards to Public and NASA People and Property Hazardous Materials Pressures, Vacuums, Temperatures
GROUND HANDLING OPERATIONS	
Accidents:	Hazards to NASA People and Property Hazardous Materials Pressures, Vacuums, Temperatures Mechanical Electrical Noise
WORKPLACE ACTIVITIES	
Accidents:	Hazards to NASA People and Property Fire, Explosion, Electrical, Mechanical, Pressures, Vacuums, Toxics, Cryogens, Suffocants, Carcinogens, Mutagens, Noise, Microwave, Laser
OPERATIONS IN SPACE	
Operational Safety:	Hazards to Crew, Systems In-Facilities Hazards EVA Hazards

TABLE 2-1. NASA ACTIVITIES AND IDENTIFIED HAZARDS
(NASA, HEADQUARTERS SAFETY DIVISION, 1987)

Risk assessment is multidisciplinary and has evolved from the three fields:

- system safety analysis
- reliability analysis
- decision analysis

Depending upon the subject and extensiveness of risk assessment, it may require supporting analyses from such scientific and engineering areas as:

- safety
- fire science
- psychology
- human factors
- explosion technology
- operations research
- statistics
- engineering — aerospace, structural, chemical, nuclear, electrical, mechanical, transportation, etc.
- geology
- climatology
- meteorology
- toxicology
- economic analysis
- environmental impact analysis
- criminology

The following U.S. Government agencies and commercial industries have experience in using risk assessment techniques:

- Nuclear Regulatory Commission and the nuclear power industry
- Chemical processing industry
- Department of Transportation
  - Hazardous materials transportation
  - Commercial and general aviation

- Environmental Protection Agency
- Department of Defense
- NASA

A brief overview of the elements of system safety analysis which support risk assessment is next given. The steps that are followed in risk assessment are then described. For additional details, reference should be made particularly to the NASA System Safety Handbook, NHB 1700.1, V.7, and the Nuclear Regulatory Commission's Probabilistic Risk Assessment Procedures Guide, NUREG/CR-2300.

2.2 SYSTEM SAFETY ANALYSIS

System safety analysis evolved in the aerospace industry in the late 1950's and early 1960's. It incorporates the following three basic procedures:

- hazards identification, via preliminary hazards analysis (PHA) and other similar techniques
- reliability analysis, such as failure modes and effects analysis (FMEA)
- logic model analysis, such as fault tree analysis (FTA).

2.3 PRELIMINARY HAZARDS ANALYSIS

The first procedure generally carried out in system safety analysis is preliminary hazards analysis (PHA). The objectives of a PHA are to identify hazardous conditions inherent in a system and to determine the significance of any potential accidents. One goal of a PHA is to prevent the recurrence of accidents that have been observed in similar activities in the past. Possible steps for PHA include the following:

Step one — Define the system for purposes of analysis, i.e., specify:

1) functional purpose of the system

   a) tasks
   b) time periods involved
   c) environmental conditions

2) component identification

   a) subsystems
   b) components

3) functional order of the system

   a) interrelationships among components and subsystems
   b) information flow within the system (inputs, outputs, logic, etc.)

Step two — Identify hazards that may exist within the system, and which may give rise to, or exacerbate, accidents.

Possible formats for this identification include:

1) narrative description,
2) tabular or matrix
3) top level fault tree.

Methods for identifying hazards that may exist within a system include:

1) checklists
2) related experience
3) engineering judgment.

A listing of representative hazards in NASA activities was given in Table 2-1.

Step three — Identify hazards that are causative factors for a potential accident

Step four — Identify hazards associated with the potential effects of accidents

Step five — Decide if hazards warrant corrective measures by considering the frequencies and severities of their associated potential accidents.

Step six — Identify hazardous events (e.g., component failures) requiring detailed "bottom-up" analysis (e.g., failure modes and effects analysis), or accident events that need to be developed by "top-down" analysis (e.g., fault tree analysis).

2.4 FAILURE MODES AND EFFECTS ANALYSIS (FMEA)

FMEA is a reliability analysis procedure that provides essential inputs to system safety analysis and risk assessment. It uses a bottom-up approach to identifying and evaluating hazards. It analyzes component failure modes and determines their possible effects on the system. Four basic hardware failure modes are:

- premature operation of a component
- failure of a component to operate at a prescribed time
- failure of a component to cease operation at a prescribed time
- failure of a component during operation.

The above failure modes may be categorized by two types of subsystem functional faults. The second and fourth failure modes represent a system element failing to perform an intended function. The first and third failure modes represent a system element performing an inadvertent function. As described in Section 3.1, below, the distinction between the two types of subsystem functional faults is important when constructing risk models such as fault trees. For example, there are sensors in the main engines on the Space Shuttle that will initiate an engine shutdown if a redline condition exists. Redline conditions are conditions such as high engine temperature or low chamber pressure. These sensors can fail in two ways: they can cause unnecessary engine shutdown due to a spurious reading or they can fail to detect a redline condition with the possibility of a catastrophic engine loss. In the first case, an inadvertent function is achieved; in the second case, an intended function is not achieved.

2.5 FAULT TREE ANALYSIS

Various other standard system safety analyses may be conducted to extend a PHA. A comprehensive alternative is fault tree analysis (FTA). FTA uses a top-down approach to identifying and evaluating hazards. It is a formalized process for delineating the possible modes of occurrence of a specified undesired event (i.e., the possible sequences of subsidiary events, or "scenarios," which can lead to the specified event) in a given system.

In system safety analysis, the fault tree is a graphic model that reveals those parallel and sequential combinations of "basic events" corresponding to component states that can result in the occurrence of a specified undesirable system state associated with a particular system-level failure or accident event. This failure or accident event is called the Top Event in the fault tree. The steps in fault tree analysis include:

- acquire an understanding of the system
- define the undesired Top Event
- construct the fault tree
- carry out a qualitative fault tree analysis
    - — find the Top Event's "minimal cutsets" (or mincutsets; see below)
    - — conduct a common cause analysis (i.e., a special analysis of non-independent component failure events)
    - — conduct a qualitative importance analysis to identify qualitatively important failures as possible subjects for improvements
- carry out a quantitative or probabilistic fault tree analysis, when appropriate
    - — compute Top Event probability or frequency (which is the risk of occurrence of the Top Event's possible consequences)
    - — compute probabilistic importance of the basic events and mincutsets to identify more precisely failures that would be subjects for improvements with the greatest payoffs
- conduct tradeoff studies of candidate improvements (or risk mitigations)
- develop results and make decisions on accepting or mitigating the Top Event risk.

FTA has a fundamental advantage over FMEA in that FTA can recognize and analyze multiple failures in an efficient manner. However, the information from an FMEA is required at the basic event level (i.e., the limit of resolution) in the fault tree, where component failure modes are expressed, and can aid the definition of the fault tree linkages at the levels just above.

See the Fault Tree Handbook (Vesely et al., 1981) for further information on fault trees, Top Events, basic events, mincutsets, etc.

2.6 STEPS IN A COMPLETE QUALITATIVE AND QUANTITATIVE RISK ASSESSMENT EMPLOYING LOGIC MODELS

The Reactor Safety Study (Nuclear Regulatory Commission, 1975) used logic models consisting of fault trees in conjunction with event trees (see below) to define reactor accident scenarios. These scenarios were generated and analyzed by a series of steps common to probabilistic risk assessment:

Step one — Identification of undesired events
Step two — Understanding of the system
Step three — Generation of the logic model
Step four — Qualitative evaluation of the logic model
Step five — Acquisition and analysis of the basic event data
Step six — Quantitative or probabilistic evaluation of the logic model
Step seven — Conduct of the sensitivity or importance analysis
Step eight — Conduct of the consequence analysis
Step nine — Conduct of the uncertainty analysis
Step ten — Peer review

Depending upon its scope and extent, all or some of the above steps are carried out in any risk assessment. In selecting the steps to be followed and the depth with which this is to be done, it is necessary to bound and scope the assessment carefully, considering the constraints on time and resources.

Steps one and two listed above are also conducted in a system safety analysis and in reliability analysis such as FMEA, as was described in Sections 2.2 and 2.4, and sometimes also in statistical analyses of tests, simulations, and observed incidents and trends. In these steps, one is concerned with identifying failures or external events that can lead to consequences such as fire, explosion, and toxic release and then losses such as injuries or deaths, or partial or total mission degradations.

Step three, logic model generation, discussed in Section 4, following, describes the generation of accident scenarios using logic models that include fault trees, event trees, and fault trees in combination with event trees. As described in Section 4.2, below, event trees are particularly useful for displaying complex time dependencies and for describing dependencies of events. In addition, event trees provide a convenient mechanism for human reliability modeling.

Other possible approaches for accident scenario modeling include GO diagrams (Williams et al., 1978; Gateley et al., 1980), cause consequence diagrams (see, e.g., Nielsen, 1975), and digraphs (see e.g., Lapp and Powers, 1977 a and b). See Sections 4.3, 4.4, and 4.5 below.

Step four, qualitative evaluation, involves the generation of the system modes of failure called the mincutsets, and the conducting of a common cause analysis (see, e.g., Rasmusson et al., 1979). See also Section 4.7, below.

Step five, data analysis, entails generating reliability data for component failure modes, for human errors, and for environmental conditions. This step may also include (in maintained systems' analyses) defining maintenance policies to which the components are subjected. See Section 4.9, below.

In step six, logic models* are employed to compute the frequencies of the accident scenarios generated in step three. Step six requires as input the reliability data generated in step five. In addition, assumptions must be made about dependencies of events. It is essential also to distinguish between two types of events in the computation of accident frequency, initiating events and enabling events.

An initiating event is an event that causes a perturbation in a system parameter which can cause the Top Event to occur. Initiating events are always defined with a Top Event in mind. Enabling events are events that permit the Top Event to occur when an initiating event occurs. The initiating event can occur before or after the enabling event. The distinction of the two types of events is important where there are mincutsets (i.e., the possible particular sequences of events which form system modes of failure) which include more than one basic event; i.e., mincutsets of order two or higher. For maintained systems, inspection intervals are important for enabling events. The longer the inspection

* Other forms of models, including direct statistical inferences, are sometimes also applicable when adequate system-level experience data are available (see Volume II of the NASA Safety Risk Management Program Plan). In most NASA applications, however, only logic tree models are likely to be used, in order to overcome the lack of system-level data.

interval, the greater the probability that a component can fail upon a demand created by the occurrence of the initiating event. It is to be noted that enabling failures (e.g., of safety systems) can exist prior to occurrence of the demand. The consideration of fault duration times of enabling events is essential.

In step seven, an importance analysis is conducted which involves combining information that is both qualitative and probabilistic in nature. One purpose of an importance analysis is to generate a ranking to determine the system and/or component failure modes that dominate the Top Event occurrence probability or risk. Such a ranking can suggest where hardware, software, human factors, and component design changes can be implemented to improve safety and/or reliability.

Also, in step seven, the effects of the various assumptions of the analysis are tested, such as on operator recovery capabilities, the sensitivities of the reliabilities of components to environmental conditions, etc.

In step eight, the consequences associated with accident scenarios are considered (see also Section 5, below). The concern is with modeling the phenomenology of the accident by considering such factors as:

- peak overpressure in an explosion
- blast effects and missile generation
- structural failures
- thermal radiation due to fire or explosion
- release rate of toxic or corrosive vapors, flammable gas, etc.
- atmospheric dispersion and weather.

By considering the above factors in the modeling of immediate physical and longer-term health effects on given population distributions, the accident consequences, such as the expected or worst-case numbers of injuries, deaths, or economic losses, can be estimated. The probabilities of the occurrence of such various possible consequences are the risks that were to be assessed.

In step nine, an uncertainty analysis is performed. It entails generating confidence intervals for accident frequencies, for consequences and losses, and, finally for risks. Uncertainty arises due to inaccuracy of the reliability data (step five), modeling uncertainty (step three), assumptions (step 7), and in the consequence calculations (step eight).

In step ten, a peer review is conducted. Ideally, the peer review is carried out by trained and experienced professionals who do not have a vested interest in the outcome, the results, or the implications of the PRA, so that the review can be fair and objective. A case in point is the peer review of the Reactor Safety Study (Nuclear Regulatory Commission, 1975) conducted by the Lewis Committee (Lewis et al., 1978).

CHAPTER 3
SYSTEM OR FACILITY DESCRIPTION AND EVENTS IDENTIFICATION

In this section is described the process of defining the undesired system- or facility-level events for purposes of risk assessment.

3.1 SYSTEM DESCRIPTION

System functions of concern to both reliability and safety are first defined. A description of the basic system configuration should be provided and supported by one-line diagrams depicting major components of the system. Physical dimensions, elevations, volumes, etc. are also included if important to the reliability of the system's operation. The supporting systems required (e.g., pneumatic system, hydraulic system, electrical power) are identified and described. The impacts of failures of supporting systems are delineated. The instrumentation available to monitor the performance of the system is identified and described. Any control logic associated with components in the system is also described. Information is provided concerning:

- system actuation: the parameters and setpoints used for automatic system actuation
- component trips: the parameters and setpoints used to automatically prevent component operation
- system isolation: the parameters and setpoints used to isolate the system.

The general schedule for system tests and changes in system configuration during these tests is described. For a maintained system, the maintenance schedule and procedures with respect to availability of system components is described. A diagram illustrating the system configuration during maintenance is provided.

If the system is manned, the roles of the operators in system performance, including manual actuation or control capabilities, is summarized. The tasks of the operators, etc., are identified. Recovery actions available to the operators are discussed for major component or system failure modes. The emergency operating procedures for each system are summarized.

The responses of the system to important accident conditions are delineated, focusing on:

- performance requirements on the system in response to postulated accident conditions (success/failure criteria for accident-caused demands on the system)
- the physical impact of accident conditions on the ability of the system to perform its function
- the impacts of the system's failure on other important systems.

Relevant operating experience should be provided. In addition, information from any previous or associated assessments (such as FMEAs) should be incorporated.

3.2 IDENTIFICATION OF UNDESIRED EVENTS

The identification of undesired events (Top Events) in the system's operation can be accomplished in numerous ways. Failures of the system to perform its intended function (reliability failures) and events that result in loss of life, injuries, or loss of mission (safety failures), are potential events to be considered. As was described in Section 2.2, one application of a PHA is to identify Top Events of fault trees. For systems that have multiple system functions, event trees can be used to define combinations of successes and failures of system functions which lead to undesired system states of varying consequences.

3.3 SIMPLIFIED EXAMPLE

A tutorial example of a pressure tank system is presented to illustrate risk assessment concepts and show how these concepts apply to NASA systems and facilities.

The system shown in Figure 3-1 discharges a flammable gas from a reservoir into a pressure tank. The pumping cycle is initiated by an operator who manually resets the timer, the timer contacts close, and the pump starts. The manual switch is normally closed. Later (well before any overpressure can exist) the timer times out and the timer contacts open. Current is denied to the pump and pumping ceases. If the timer contacts do not open, the operator is instructed to observe the pressure gauge and to open the manual switch, thus causing pumping to cease. After each cycle, the compressed gas is discharged by opening the valve and then closing the valve before the next cycle begins. At the end of

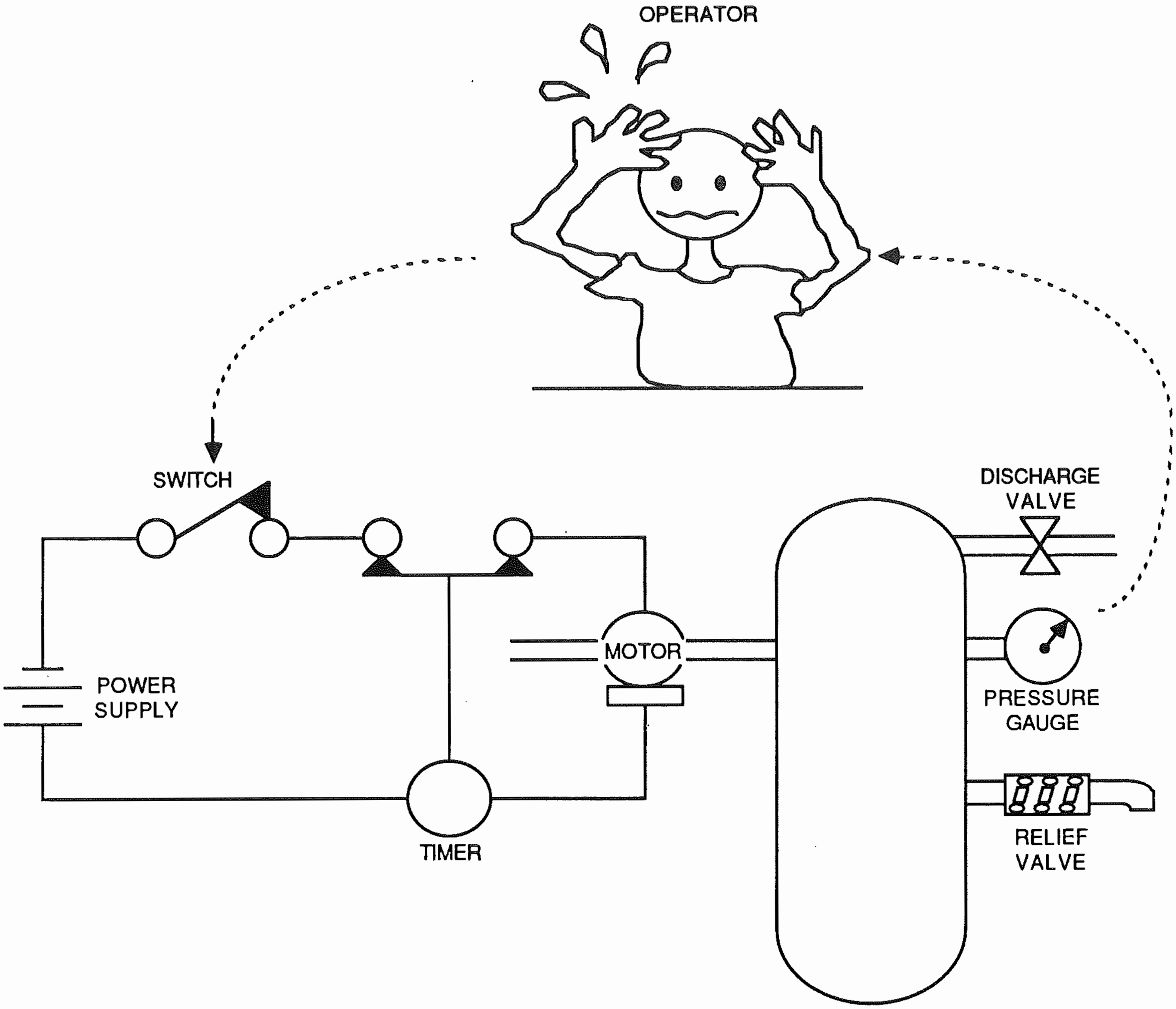


FIGURE 3-1. PRESSURE TANK SYSTEM

an operating cycle, the operator is instructed to verify the operability of the pressure gauge by observing a decrease in the tank pressure as the discharge valve is opened. Each cycle consists of a pressurization and depressurization time period. It is assumed that each cycle takes on the average one hour and that time for depressurization is negligible compared to pressurization.

The pressure tank system is enclosed in a room with a nitrogen purge system. Each day the operator is instructed to observe the detectors indicating nitrogen purge pressure and flammable gas concentration in the room. If either the nitrogen purge pressure is low or the flammable gas concentration is high (25 percent within the lower explosive limit, or LEL), the operator is instructed to shut the system down by opening the manual switch and by opening the discharge valve to depressurize the system. Then he is instructed to do any necessary repairs for restoring the nitrogen purge system or correcting leaks or malfunctions within the pressure tank system.

3.3.1 System Functions

The pressure tank system performs pressurization and depressurization functions that are necessary for the system to reliably perform as intended. In addition, the system executes safety functions. The operator proceeds with system shutdown if the tank pressure is too high. The relief valve causes depressurization when needed to prevent rupture of the tank. Also, there are safety functions performed to prevent fire or explosion within the room where the pressure tank system is housed. The operator shuts the system down in the event of loss of nitrogen purge or high flammable gas concentration within the room.

3.3.2 Undesired System Events

In this section are considered the undesired events in the risk assessment of the pressure tank system. As was described in Section 3.2, this entails identifying events that result in failure to perform the system's intended function (reliability failures) and events that result in loss of life, injuries, or loss of the mission (safety failures). Undesired events for the pressure tank system include:

- failure to pressurize tank (reliability)
- failure to depressurize tank (reliability)
- pressure tank rupture (safety)

- fire or explosion (safety)
- asphyxiation (safety)
- burns and trauma (safety).

In defining undesired events, it is also important to define associated success criteria. Dual logic applied to success criteria defines the failure criteria used in a risk model such as a fault tree. For example, a successful launch of the Space Shuttle requires all three engines to be functioning immediately after launch (a one-out-of-three failure criterion). However, with two engines working and one having failed immediately after launch, the Space Shuttle can successfully abort and land without a catastrophic loss of the vehicle (a two-out-of-three failure criterion).

CHAPTER 4
LOGIC MODELS FOR QUALITATIVE AND QUANTITATIVE RISK ASSESSMENT — SYNTHESIS AND EVALUATION

In this section, five topics are discussed:

- logic model generation
- qualitative evaluation of the logic model
- quantitative or probabilistic evaluation of the logic model
- sensitivity and uncertainty analysis
- evaluation of mitigative measures

Each of the following five logic models used in risk assessment is described, with its advantages and disadvantages, particularly for applications when rare events are of concern and direct experience with system failures is lacking:

- fault trees
- event trees (in combination with fault trees)
- GO methodology
- cause-consequence diagrams
- directed graphs.

The qualitative evaluation of each logic model includes:

- identifying initiating and enabling events
- generation of mincutsets
- conducting of a common cause analysis
- computation of structural importance.

Structural importance is a qualitative measure of the importance of a component to system operation. For example, when considering the failure of a system to function as intended, a component placed in series with the remainder of the system is more important structurally to the functioning of the system than that same component placed in parallel with the remainder of the system.

The quantitative evaluation of a logic model generally requires its Boolean representation, such as its group of mincutsets, together with knowledge of maintenance policies, as applicable, and basic event reliability data such as on-demand and continuous operation component failure rates, and possibly including human operator failure probabilities. Repair times and inspection intervals must be given for analyses of maintained systems. Accident frequency estimates and probabilistic evaluations of importance for basic events and mincutsets are then derived.

In this section, also, a comparative evaluation of the five logic models is presented. Their use in sensitivity analysis, in the evaluation of risk mitigation measures, and in uncertainty analysis, is discussed.

Finally, available computer codes are identified.

4.1 FAULT TREES

For a comprehensive description of conventional fault tree analysis, refer to the Nuclear Regulatory Commission's Fault Tree Handbook (Vesely et al.,1981).

Fault tree construction consists of three basic steps:

- defining the system to be analyzed and the boundaries of the system

- defining the top undesired event in the fault tree

- constructing the fault tree to the limit of resolution using the fault tree construction rules given in Table 4-1.

Table 4-1 outlines the construction of fault trees according to the immediate cause principle, i.e., from sequences of causally-related steps. It is important to keep in mind that if the Top Event changes, the fault tree logic and hence the subsequent evaluation of the fault tree also must change.

TABLE 4-1. FAULT TREE CONSTRUCTION RULES

Rule 1: State the fault event as a fault, including the description and timing of a fault condition at some particular time. Include:

a. What the fault state of that system or component is.
b. When that system or component is in the fault state.

Test the fault event by asking:

c. Is it a fault?
d. Is the what-and-when portion included in the fault statement?

Rule 2: There are two basic types of fault statements: state-of-system and state-of-component.

a. If the fault statement is a state-of-system statement, use Rule 3.
b. If the fault statement is a state-of-component statement, use Rule 4.

Rule 3: A state-of-system fault may use an AND, OR, inhibit gate, or no gate at all. To determine which gate to use, the faults must be the:

a. Minimum necessary and sufficient fault events.
b. Immediate fault events.

Rule 4: A state-of-component fault always uses an OR gate. To continue, look for the primary, secondary, and command failure fault events. Then state those fault events.

a. Primary failures are failures of that component within the design envelope or environment.
b. Secondary failure are failures of that component due to excessive environments exceeding the design environment.
c. Command faults are inadvertent operations of the component because of failures of control elements.

Rule 5: No gate-to-gate relationships.

Rule 6: Expect no miracles; those things that would normally occur as the result of a fault will occur, and only those things. Also, normal system operation may be expected to occur when fault occurs.

Rule 7: In an OR gate, if any input exists, the output exists. Fault events under the gate may be restatements of the output events.

Rule 8: An AND gate defines a causal relationship. If the input events coexist, the output is produced.

Rule 9: An inhibit gate describes a causal relationship between one fault and another, but the indicated condition must be present. The fault is the direct and sole cause of the output, when that specified condition is present. Inhibit conditions may be faults <u>or</u> situations, which is why AND and inhibit gates differ.

The fault tree for failure to pressurize the tank in the pressure tank example is shown in Figure 4-1. The first three events below the top level OR gate are events that result in loss of pressure. The fourth event, "pump motor fails to operate," corresponds to failure of the system to pressurize. It is seen that the fault tree in Figure 4-1 contains all OR gates, as is common for reliability-type failures.

The fault tree for pressure tank rupture under load or due to overpressure is shown in Figure 4-2. For this fault tree, the assumption is made that the pressure tank system starts each cycle unpressurized. Assumptions are important in any risk assessment and they should be listed clearly.

The salient features of the fault tree in Figure 4-2 are now examined. The fault tree consists of gate events and basic or primary events. Gate events are outputs of logic gates, either AND or OR; basic events appear at the bottom of the fault tree and represent the limit of resolution of the fault tree. Basic events include:

- random equipment failures
- human errors
- environmental conditions.

Basic events can include common cause events such as failures in support systems. The event, "pressure tank rupture under normal load," is a single event leading to rupture of the tank. This event represents a passive failure. Consider the gate event, "tank rupture due to overpressure." The cause of overpressure is the gate event, "timer contacts fail to open," which causes the pump motor to continue to operate (i.e., a component fails to cease operation at a prescribed time). The basic event, "pressure relief valve fails to operate," represents failure of pressure protection when the pump motor continues to operate (i.e., a component fails during operation). The gate event, "current through manual switch contacts too long," represents failure of the operator shutdown function (an inadvertent function is achieved). The basic event, "voltage surge," is a common cause initiating event (also referred to as a special initiator); i.e., it is an event that causes a system upset condition and simultaneously fails system mitigative features. Failures in support systems such as electric power can be common cause initiating events. External events, such as floods, fires, or earthquakes, can also be common cause initiating events.

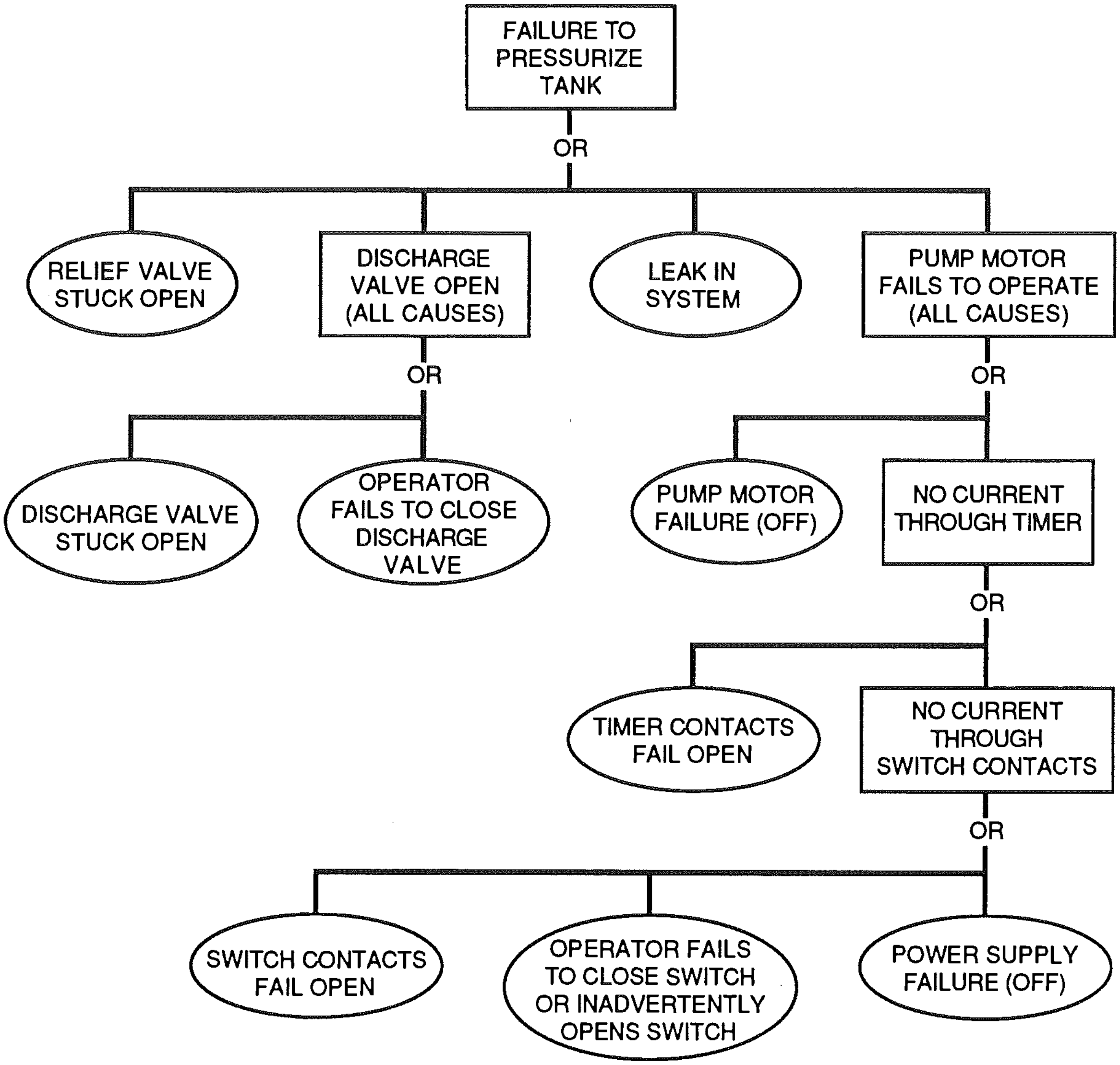

FAILURE TO PRESSURIZE TANK
OR
RELIEF VALVE STUCK OPEN
DISCHARGE VALVE OPEN (ALL CAUSES)
LEAK IN SYSTEM
PUMP MOTOR FAILS TO OPERATE (ALL CAUSES)
OR
DISCHARGE VALVE STUCK OPEN
OPERATOR FAILS TO CLOSE DISCHARGE VALVE
OR
PUMP MOTOR FAILURE (OFF)
NO CURRENT THROUGH TIMER
OR
TIMER CONTACTS FAIL OPEN
NO CURRENT THROUGH SWITCH CONTACTS
OR
SWITCH CONTACTS FAIL OPEN
OPERATOR FAILS TO CLOSE SWITCH OR INADVERTENTLY OPENS SWITCH
POWER SUPPLY FAILURE (OFF)

FIGURE 4-1. FAULT TREE FOR FAILURE TO PRESSURIZE TANK

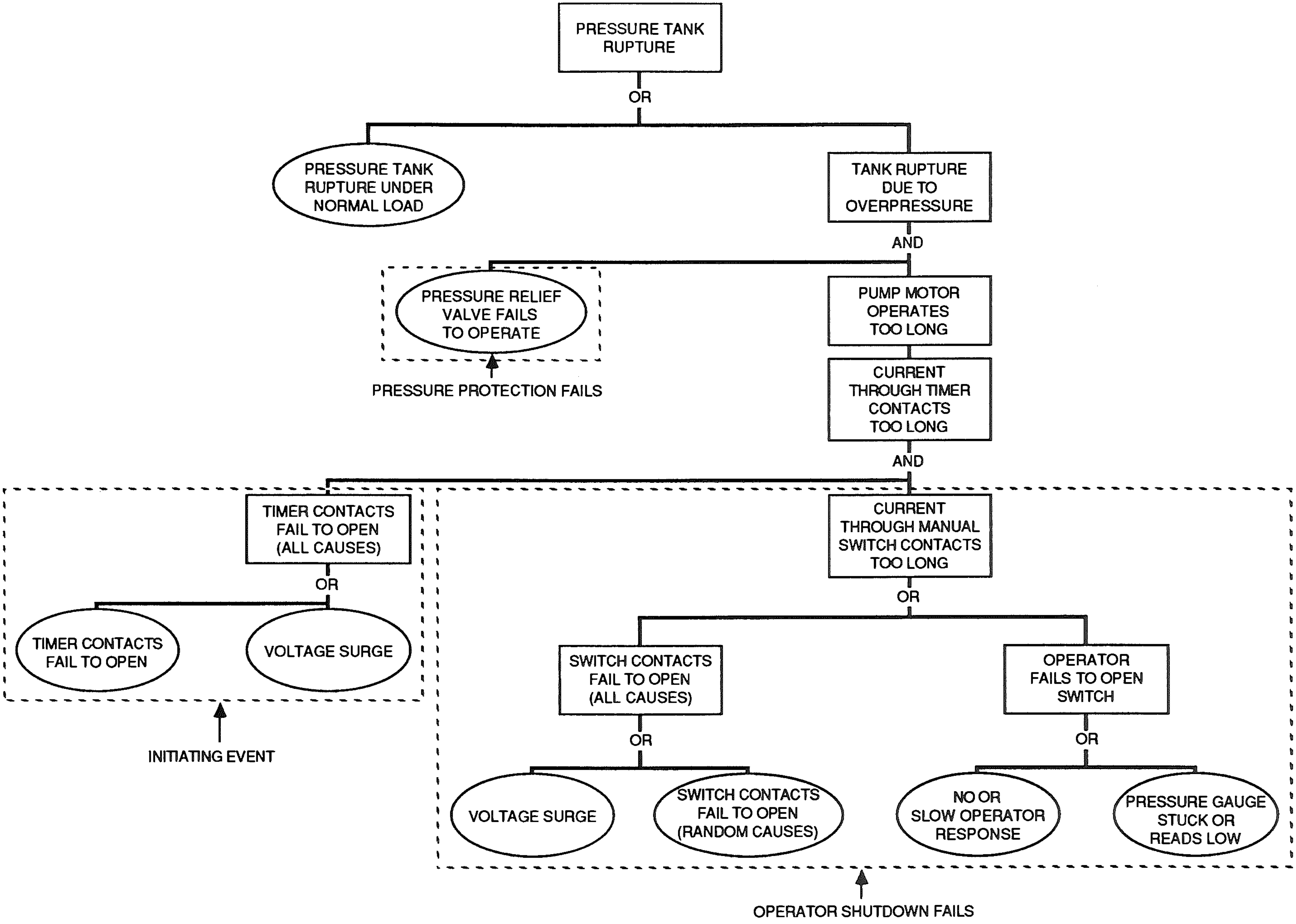

PRESSURE TANK RUPTURE
OR
PRESSURE TANK RUPTURE UNDER NORMAL LOAD
TANK RUPTURE DUE TO OVERPRESSURE
AND
PRESSURE RELIEF VALVE FAILS TO OPERATE
PRESSURE PROTECTION FAILS
PUMP MOTOR OPERATES TOO LONG
CURRENT THROUGH TIMER CONTACTS TOO LONG
AND
TIMER CONTACTS FAIL TO OPEN (ALL CAUSES)
OR
TIMER CONTACTS FAIL TO OPEN
VOLTAGE SURGE
INITIATING EVENT
CURRENT THROUGH MANUAL SWITCH CONTACTS TOO LONG
OR
SWITCH CONTACTS FAIL TO OPEN (ALL CAUSES)
OR
VOLTAGE SURGE
SWITCH CONTACTS FAIL TO OPEN (RANDOM CAUSES)
OPERATOR FAILS TO OPEN SWITCH
OR
NO OR SLOW OPERATOR RESPONSE
PRESSURE GAUGE STUCK OR READS LOW
OPERATOR SHUTDOWN FAILS

FIGURE 4-2. FAULT TREE FOR PRESSURE TANK RUPTURE

The fault tree for fire and explosion is given in Figure 4-3, consisting of two sheets. For fire or explosion to occur, the following conditions (assuming reactants are below the autoignition temperature) are needed:

(1) heat or ignition source present
(2) flammable species present between the lower and upper flammability limits
(3) oxygen present above minimum concentration for combustion.

Fire or explosion can result when all of these three conditions occur, in any time sequence.

4.2 EVENT TREES

An event tree is a logic diagram that starts with an initiating event and defines the resulting possible combinations of success and failure events which lead to various outcome system states.

Figure 4-4 shows the event tree for pressure tank rupture due to overpressurization. The event tree starts with an initiating event and describes sequences of failures (SOF) of system mitigative features that can lead to undesired system or plant states. In Figure 4-4, PO denotes the event "pump overrun," the initiating event. OS denotes the failure of the operator shutdown system. PP denotes the failure of the pressure protection system. The headings on the event tree are logically ordered in time. The reason that pressure protection is considered as the last protective feature is that it is assumed that there is very little time for the operator to respond to prevent tank rupture given that the setpoint pressure has been reached and the pressure relief valve has failed to operate.

There are three sequences displayed at the terminal nodes of the event tree in Figure 4-4. The sequence labeled PO*OS*PP causes overpressure and tank rupture (* denotes logical intersection, AND). The other two sequences lead to safe results. The event tree defines Top Events of fault trees. It is seen that portions of the fault tree described in Figure 4-2 appear in Figure 4-4. Note that the event tree in Figure 4-4 contains an initiating event fault tree.

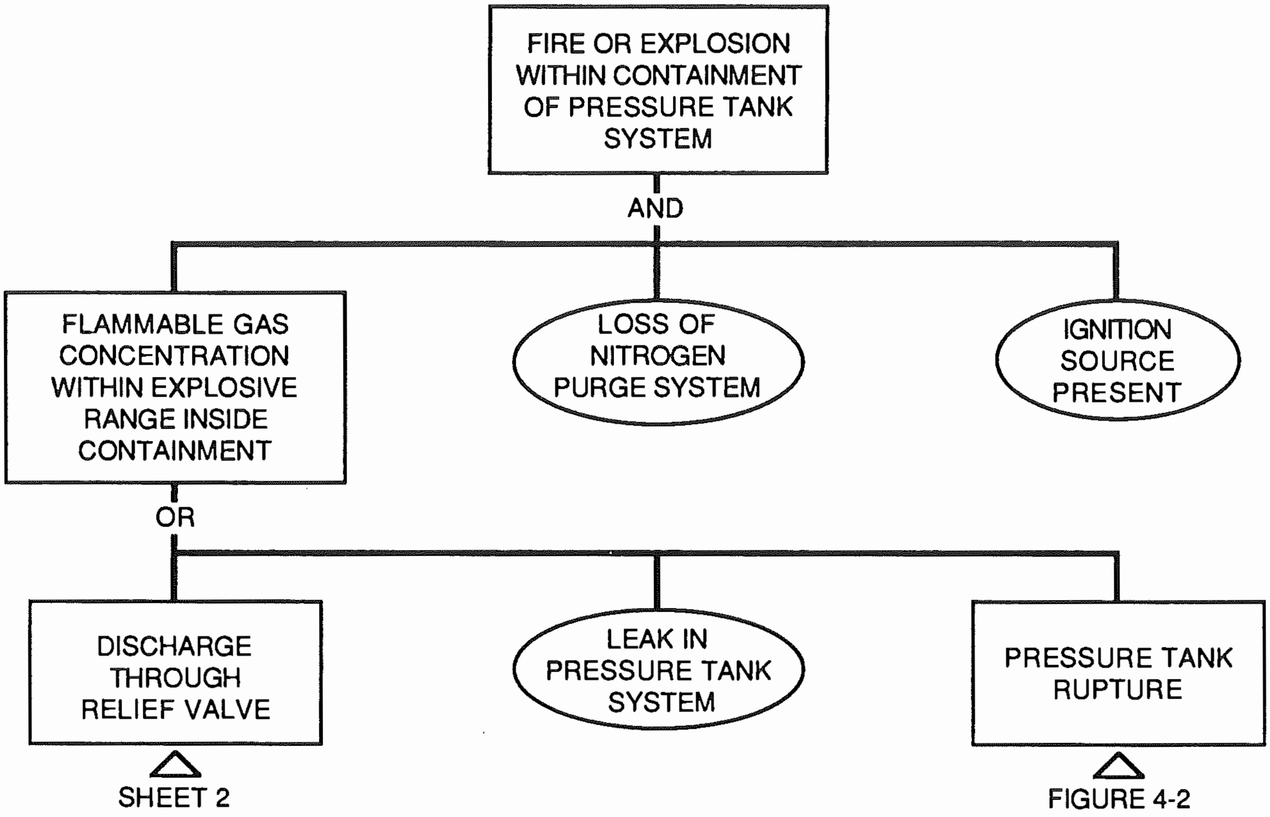

FIRE OR EXPLOSION WITHIN CONTAINMENT OF PRESSURE TANK SYSTEM
AND
FLAMMABLE GAS CONCENTRATION WITHIN EXPLOSIVE RANGE INSIDE CONTAINMENT
LOSS OF NITROGEN PURGE SYSTEM
IGNITION SOURCE PRESENT
OR
DISCHARGE THROUGH RELIEF VALVE
LEAK IN PRESSURE TANK SYSTEM
PRESSURE TANK RUPTURE
SHEET 2
FIGURE 4-2

FIGURE 4-3. FAULT TREE FOR FIRE OR EXPLOSION (SHEET 1)

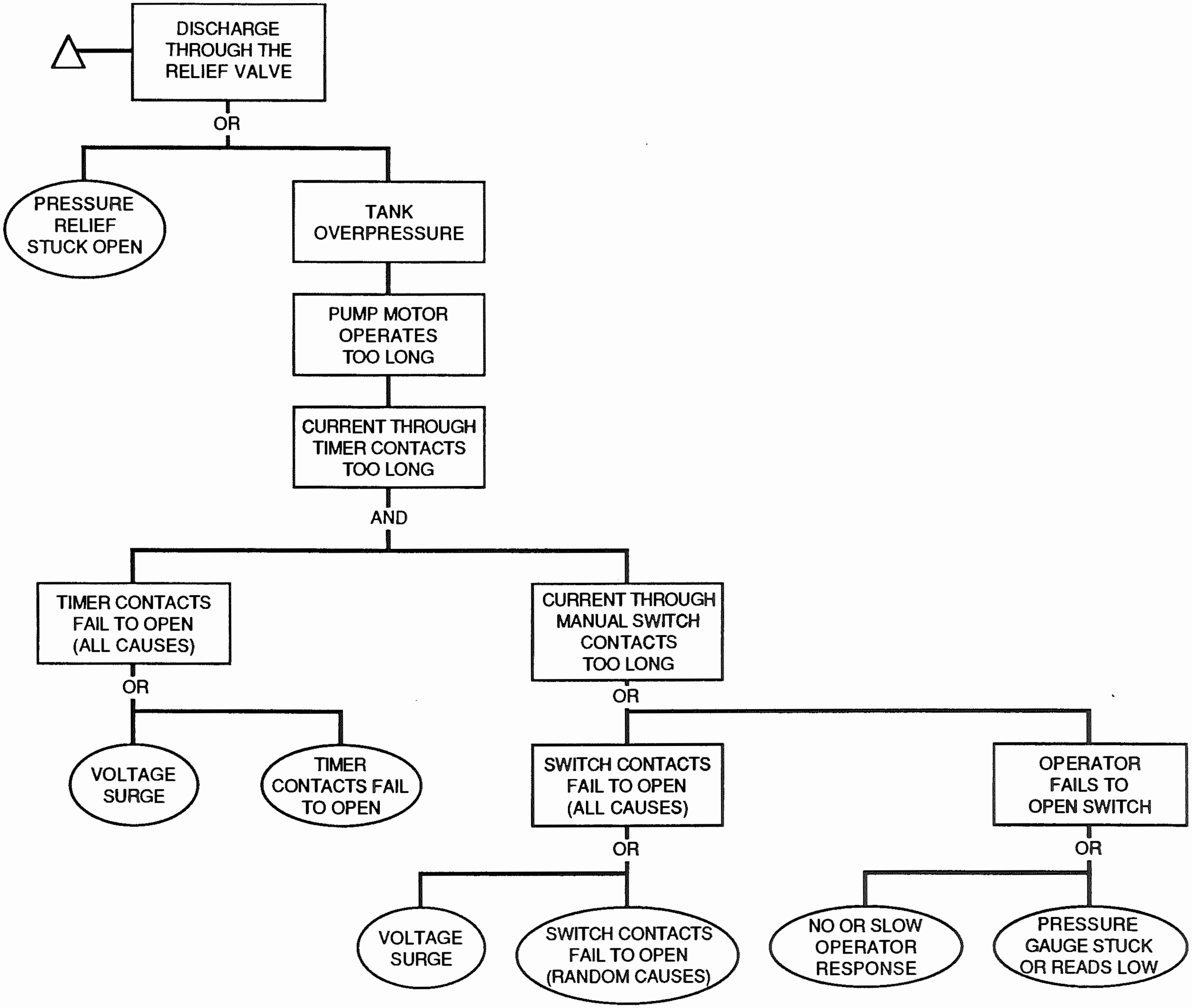

DISCHARGE THROUGH THE RELIEF VALVE
OR
PRESSURE RELIEF STUCK OPEN
TANK OVERPRESSURE
PUMP MOTOR OPERATES TOO LONG
CURRENT THROUGH TIMER CONTACTS TOO LONG
AND
TIMER CONTACTS FAIL TO OPEN (ALL CAUSES)
OR
VOLTAGE SURGE
TIMER CONTACTS FAIL TO OPEN
CURRENT THROUGH MANUAL SWITCH CONTACTS TOO LONG
OR
SWITCH CONTACTS FAIL TO OPEN (ALL CAUSES)
OR
VOLTAGE SURGE
SWITCH CONTACTS FAIL TO OPEN (RANDOM CAUSES)
OPERATOR FAILS TO OPEN SWITCH
OR
NO OR SLOW OPERATOR RESPONSE
PRESSURE GAUGE STUCK OR READS LOW

FIGURE 4-3. FAULT TREE FOR FIRE OR EXPLOSION (SHEET 2)

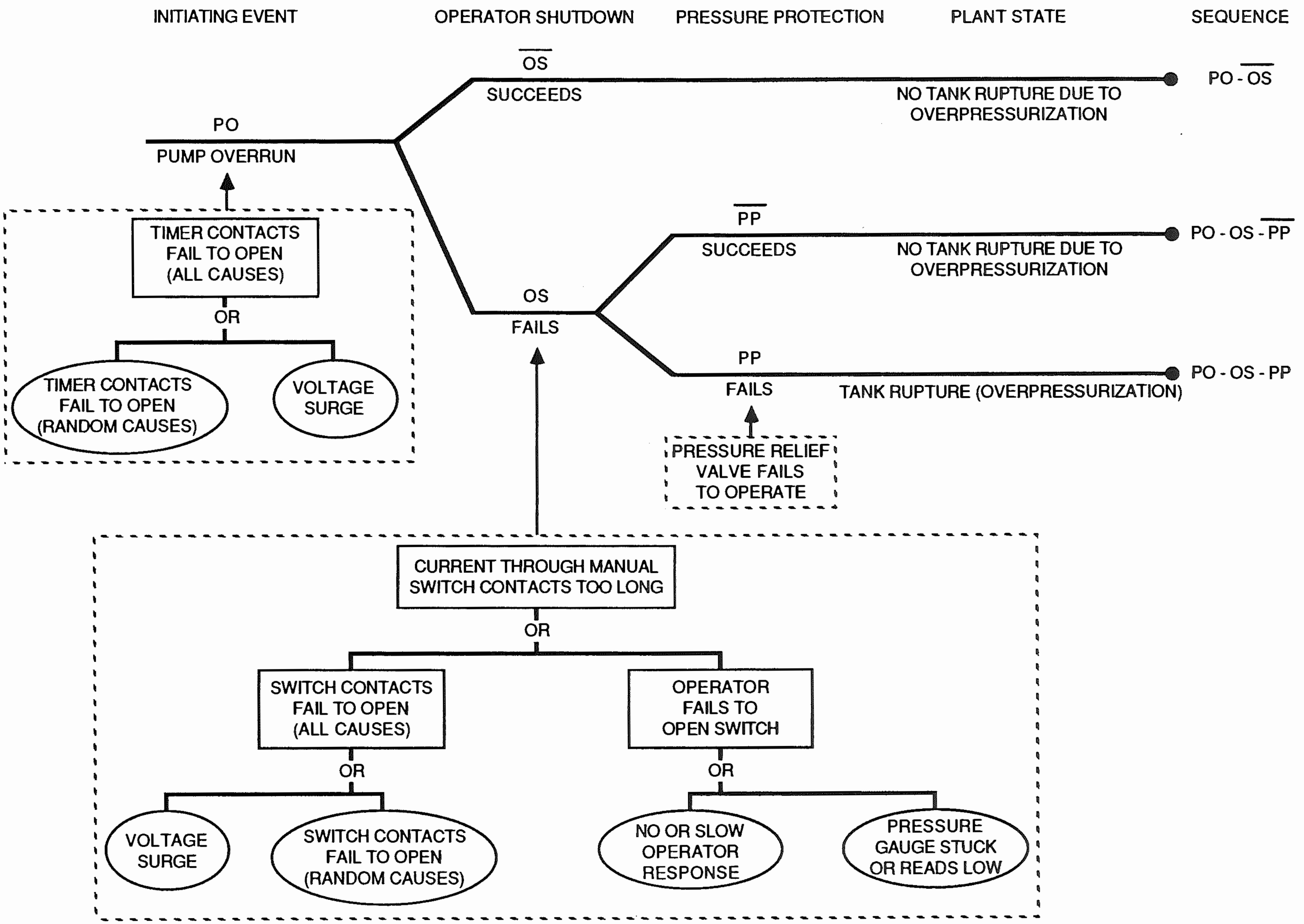

INITIATING EVENT
OPERATOR SHUTDOWN
PRESSURE PROTECTION
PLANT STATE
SEQUENCE
PO
PUMP OVERRUN
$\overline{OS}$
SUCCEEDS
NO TANK RUPTURE DUE TO OVERPRESSURIZATION
PO - $\overline{OS}$
OS
FAILS
$\overline{PP}$
SUCCEEDS
NO TANK RUPTURE DUE TO OVERPRESSURIZATION
PO - OS - $\overline{PP}$
PP
FAILS
TANK RUPTURE (OVERPRESSURIZATION)
PO - OS - PP
TIMER CONTACTS FAIL TO OPEN (ALL CAUSES)
OR
TIMER CONTACTS FAIL TO OPEN (RANDOM CAUSES)
VOLTAGE SURGE
PRESSURE RELIEF VALVE FAILS TO OPERATE
CURRENT THROUGH MANUAL SWITCH CONTACTS TOO LONG
OR
SWITCH CONTACTS FAIL TO OPEN (ALL CAUSES)
OPERATOR FAILS TO OPEN SWITCH
OR
OR
VOLTAGE SURGE
SWITCH CONTACTS FAIL TO OPEN (RANDOM CAUSES)
NO OR SLOW OPERATOR RESPONSE
PRESSURE GAUGE STUCK OR READS LOW

FIGURE 4-4. EVENT TREE FOR PRESSURE TANK SYSTEM

By identifying system functions, the equivalency of the fault tree logic in Figure 4-3 with the event tree logic in Figure 4-4 is evidenced. Event trees are especially useful for displaying functions that depend upon time. In addition, the terminal nodes on an event tree can represent outcome states with different consequences. For example, an event tree for explosion of the pressure tank system is given in Figure 4-5. The fault tree in Figure 4-3 models the initiating event for this event tree. Note that containment failure leads to the possibility of injuries or deaths as well as total system loss. System loss, but not injuries or deaths, can also occur without containment failure.

The potential for pressure tank explosion exists for Space Shuttle operations, e.g., in ground operations in which liquid hydrogen and oxygen are loaded in the external tank, and when the Space Shuttle is in flight. The exact consequences of an explosion depend upon the time at which the explosion occurs. Explosion before liftoff could result in loss of vehicle and crew as well as ground facilities (GF) and GF personnel. Thirty seconds after liftoff, an explosion would not affect GF or GF personnel. A time-phased event tree can be constructed that considers various explosion scenarios as they occur in time with the identification of associated consequences for each scenario.

Generally, both event trees and fault trees are used in a risk assessment and it depends upon the particular application whether to use only fault trees, or event trees in conjunction with fault trees. In the latter case, event trees can display complex time dependency, as in a sequence of operational phases of a system, as well as make precise the definitions of the Top Events for the fault trees associated with different time periods.

4.3 GO METHODOLOGY

The GO methodology, unlike failure-oriented fault tree analysis, is success-oriented. The GO method models system performance, as also does an event tree, in terms of system response modes, both successes and failures. The methodology evolved from its application in the defense industry and has been adapted particularly to analysis of fluid and hydraulic systems. The following description of the GO methodology is taken from Kelly and Stillwell (1981).

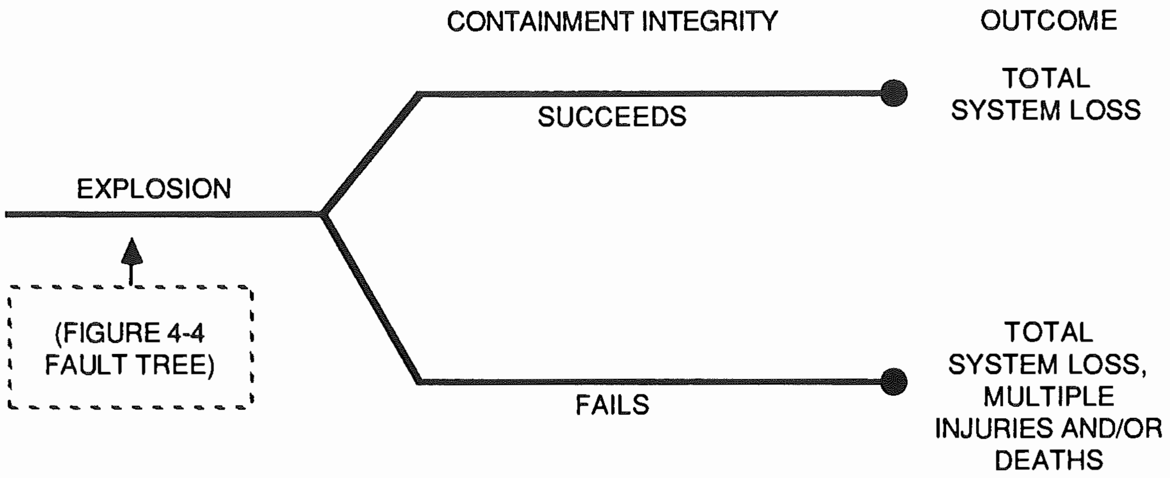


FIGURE 4-5. EVENT TREE FOR EXPLOSION OF THE PRESSURE TANK SYSTEM

4.3.1 GO Model Description

The GO methodology is a success-oriented, probabilistic combinatorial analysis procedure. Component operating probabilities and interactions are combined to produce the probabilities of desired output events. The modeling methodology includes two major elements:

- A set of standardized functional operators that are used to model physical components with mathematical entities.
- A modeling technique whose result, called a GO chart, corresponds closely to the physical layout or design schematic of the system analyzed

In the following, a brief description of the GO modeling method, including examples of each of these elements, is presented. Further details may be found in Gately et al. (1980).

4.3.2 GO Operator Types

The logical operator types combine the input event probability distributions to produce the distribution for the output event in the various states. The current set of 17 operators is shown in Figure 4-6 and is briefly described. Although 17 operators are available to the analyst, typically only five or six different operators need be employed in the solution of a given problem.

Type 1 is a simple component with one input and two operational states (good or failed). It is used extensively to model anything from a resistor to a complete subsystem. "Kind" data include P1 (good) and P2 (failed). "Kind" identifies the probability data that are assigned to the component, e.g., the probability it is good, the probability it has failed in mode 1, the probability it has failed in mode 2, etc.

Type 2 is a logical OR gate with up to 10 inputs. This type is "perfect" and is not given a kind number. The output event occurs as soon as any input appears.

Type 3 is a component with three operating modes: good (output occurs when input occurs), failed (no output), and premature (output occurs with no input). Kind data are P1 (good), P2 (failed), and P3 (premature).

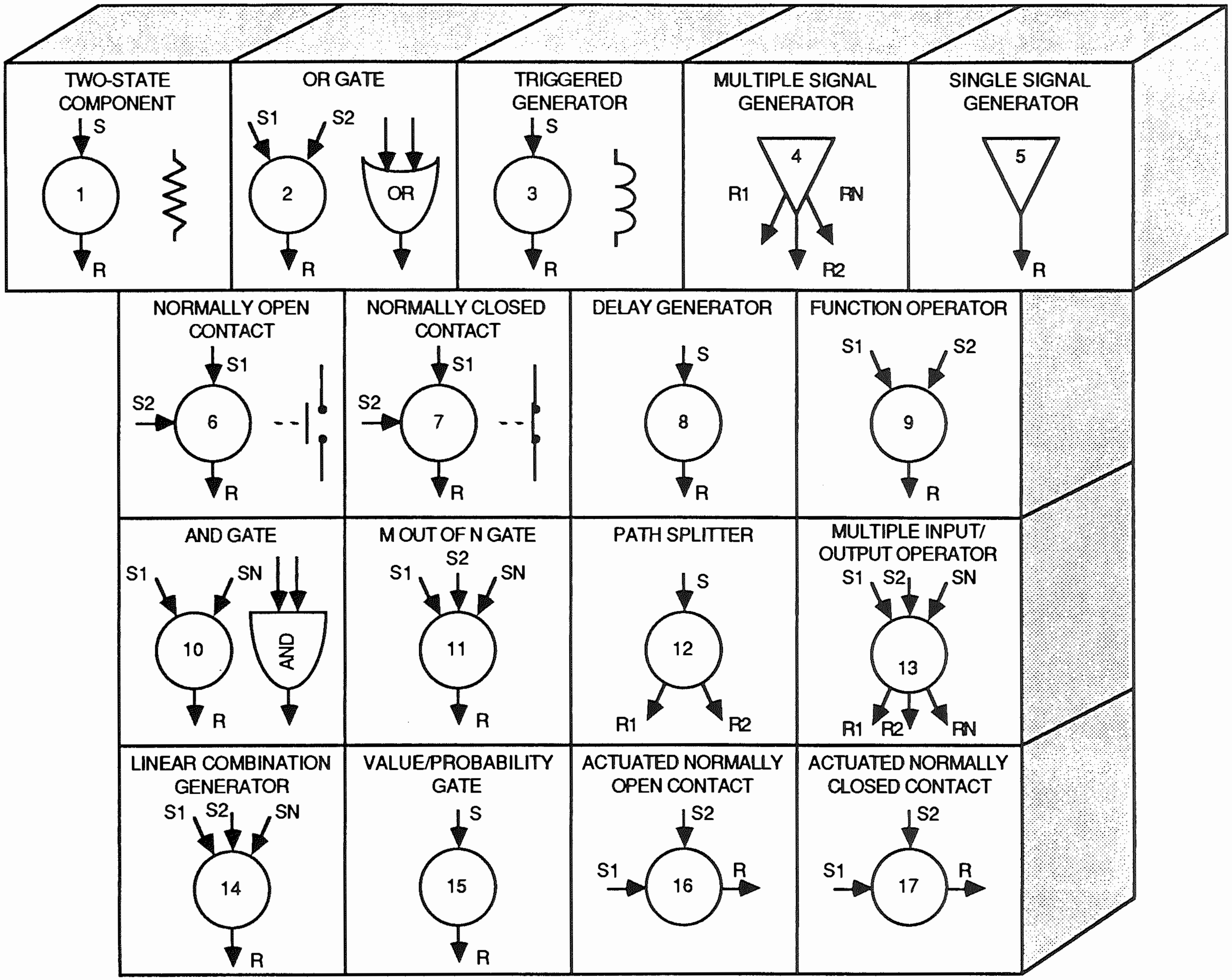


FIGURE 4-6. GO OPERATOR TYPES

Type 4 is a problem initiator that has no input of its own and which is capable of generating two or more statistically dependent signals.

Type 5 is a problem initiator or input which has no input of its own. The kind data are the probabilities of occurrence in the various time periods. This type indicates the presence, absence, or distribution of inputs (electricity, water, etc.) at the start of the problem.

Type 6 is a component with a primary input S1 and a secondary (or trigger) input S2. Output occurs when both inputs are present. The secondary input may be represented on the GO chart by a small circle or a half-arrow. A premature mode of operation is available which produces an output only when the S1 is present. The inputs are not interchangeable. Type 6 typically models a normally-open switch or a normally-closed valve. When used to model a switch, premature represents "contact shorted," and failure represents "stuck open." Kind data are P1 (good), P2 (failed), and P3 (premature).

Type 7 is similar to Type 6, but is normally closed.

Type 8 is used to model delays in component responses.

Type 9 is a general purpose, state-change operator that produces an output at a time determined by the difference between the times of S1 and S2. It can be used to model the more complex logic gates (NAND, exclusive OR, etc.). Kind data define the operator logic.

Type 10 is a perfect logic AND gate with up to 10 inputs. The output occurs only if all inputs are present.

Type 11 is a perfect logic m-out-of-n gate.

Type 12 is a disjoint path splitter. The mutually exclusive outputs are assigned probabilities and times.

Type 13 is a multiple input/output operator that provides for complete generality in tailoring output states based upon given input states.

Type 14 is a linear combination generator.

Type 15 is a value/probability gate used to control an output depending on the value of the input.

Type 16 is an actuated normally-open contact. The contact is closed at the start of the problem and opens when S2 arrives. This type is used extensively for the "interrupt" logic typical of nuclear reactor trip systems, for instance.

Type 17 is an actuated normally-closed contact. The contact is opened at the start of the problem and closed when S2 arrives.

A unique aspect of the GO operator types is that more than just binary states of a component may be represented. Types 3, 4, 6, 7, 12, 13, 16, and 17 allow three or more probability states. In the modeling process, these states are represented by the assignment of integer signal values that represent either real or artificial sequencing of events important to system function. In such a manner, multiple outcomes may be flagged by the analyst.

The GO modeling process also allows the definition of supertypes. A supertype is a structured collection of operators which the analyst chooses to treat as a single entity, usually because the collection represents a physical subsystem of the modeled system which occurs several times within the system. The supertype allows the subsystem to be modeled only once rather than at each occurrence.

4.3.3 GO Chart

The GO chart shows the logical relationships that exist between the model operators. In many cases, there is almost a one-to-one correspondence between the operators and the components of the real system.

A simple system, as shown in Figure 4-7, may be employed to illustrate the elements of a GO chart. The system consists of a common water supply tank, two pumps and associated check valves, and four level control valves (two per pump train). Auxiliaries required for system operation include power to each of the pumps and actuation signals to each level

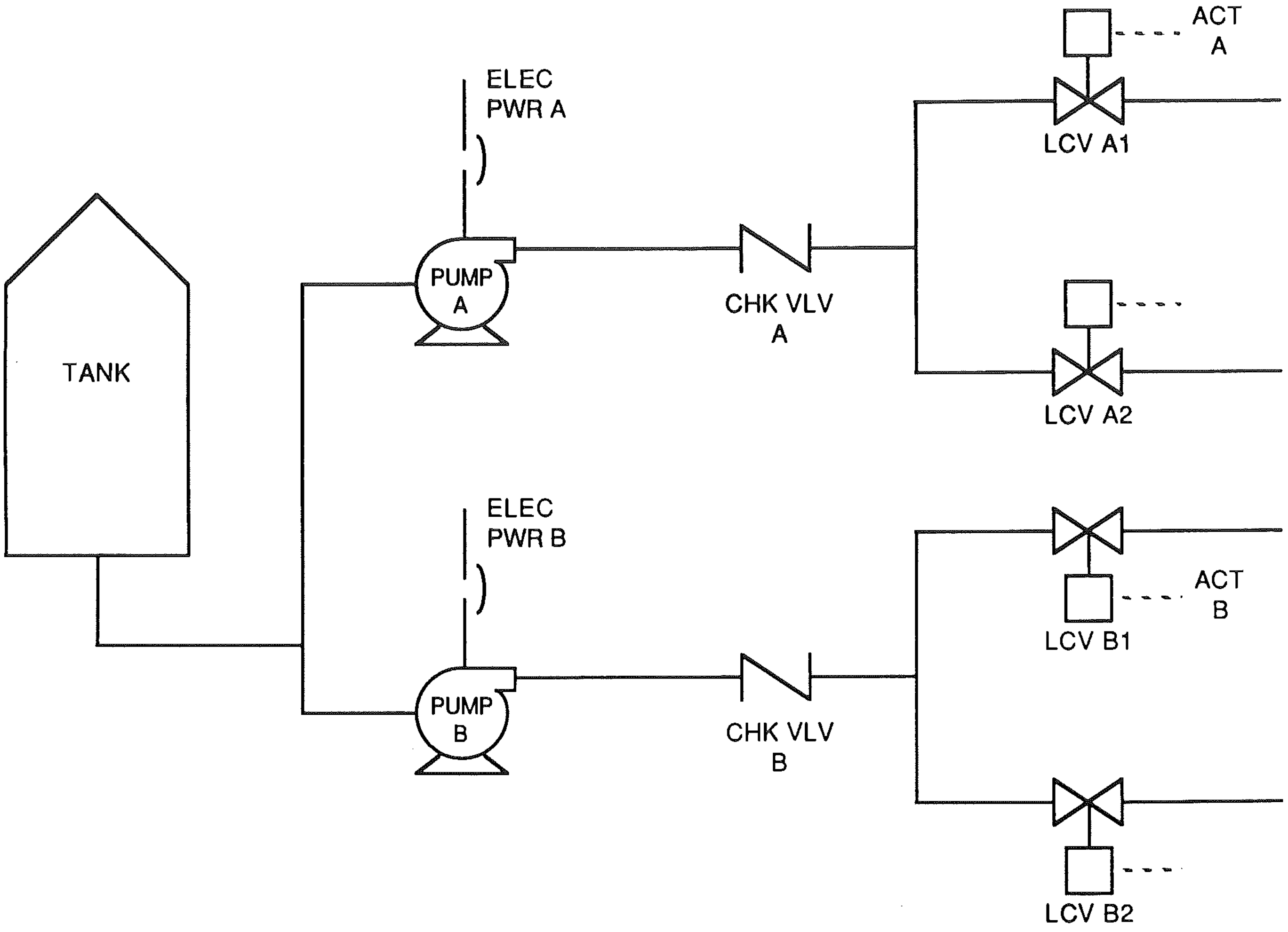


FIGURE 4-7. EXAMPLE SYSTEM DIAGRAM

control valve (LCV). Successful operation of the system involves startup of the pumps, opening of the check valves, and flow modulation by the level control valves. Appropriately modulated flow from the tank, through the pump, check valve, and two or more level control valves constitutes system success.

Figure 4-8 shows the GO chart for this example system. Each logic element of the GO chart is represented by a triangle or circle symbol (the former representing a source or initiating element). Within the logic element symbol, the leftmost number identifies the operator type (as described in the previous section); the right number identifies the operator "kind."

In the GO chart, the tank, electric power, and valve actuation elements are sources of water, power, and actuation logic, respectively, for the example system. They are therefore shown as Type 5 operators. The pumps require two inputs, water at the pump suction and electric power to drive the pump.

Therefore, Type 6 is the logical choice. The same applies to the level control valves. The check valves may logically be shown as the Type 1 GO operator. Finally, the success logic for the system is identified by the Type 11 operator in the diagram. In this case, it is assumed that flow provided by two or more of the level control valves represents system success.

The GO operators in the chart may be seen to be connected in accordance with the process flow logic of the system. Water flows from the tank through the two pumps and check valves and through the four level control valves. The lines connecting the operators are numbered by so-called "signal numbers" that describe the desired sequencing of operations. (The sequencing scheme shown in Figure 4-8 is not unique; others may be employed.) Signal 14 represents the output state of the system whose probability is to be quantified.

The assignment of kind numbers to the GO chart is straightforward. The tank as a unique component in the diagram is assigned a unique kind number. The parallel pumps, electric power sources, check valves, level control valves, and actuation signals are assumed to have identical failure probabilities so that only one kind number need be assigned to each of these redundant components. No kind number need be identified for the Type 11 since it is a "perfect" logic operator.

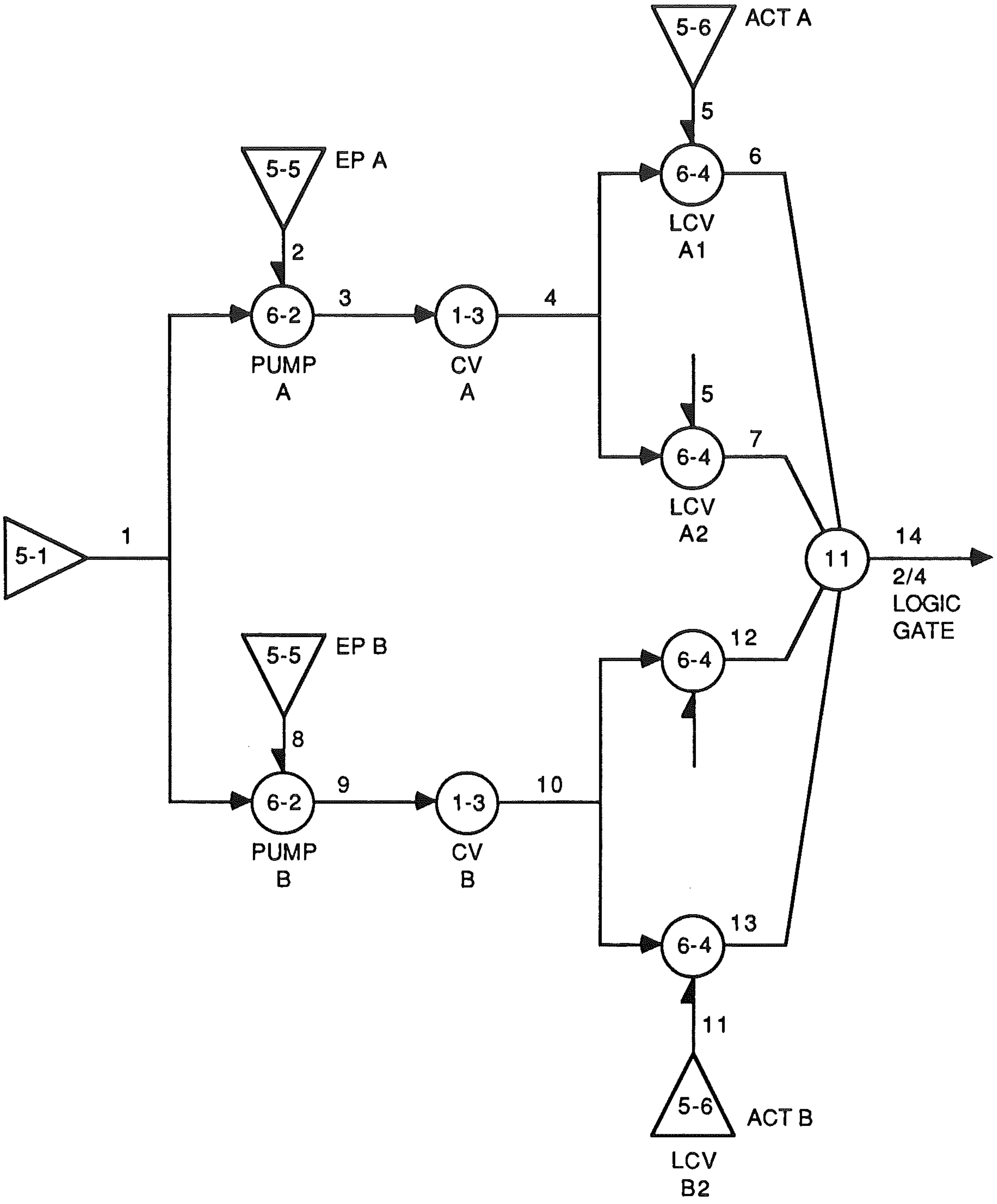


FIGURE 4-8. GO CHART OF EXAMPLE SYSTEM

By comparing Figures 4-7 and 4-8, it may be noted that the GO chart very closely simulates the actual system schematic. With the exception of the rightmost Type 11 operator in the GO chart, there is a one-to-one correlation of GO operators with components.

4.3.4 Modeling Example

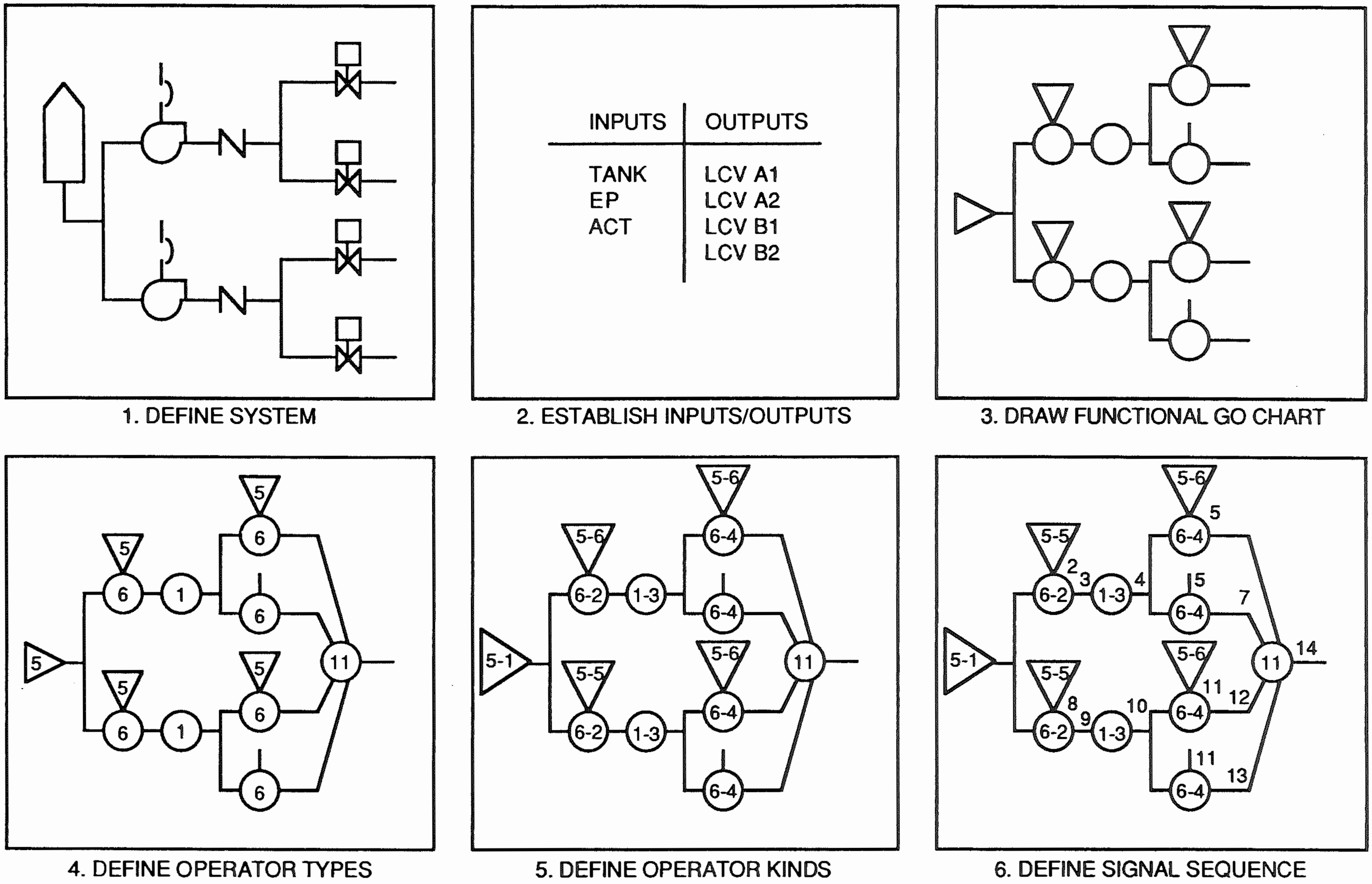
To derive the GO chart for the Sequoyah nuclear power plant's auxiliary feedwater system, as described in Kelly and Stillwell (1981), the six basic steps illustrated in Figure 4-9 were followed.

First, it was necessary to define the system sufficiently so that modeling could begin (Step 1). This involved establishing system boundaries and a desired level of detail, as well as general information gathering.

To develop the GO chart, the system inputs and outputs were first identified (Step 2). For an "open loop" system, such as the fluid system that was illustrated in Figure 4-7, this can be straightforward. For a "closed loop" system, such as an electrical network, the analyst must decide where to "break" the loop to begin the modeling process.

In Step 3, the functional GO chart is drawn showing the logical connection of system hardware starting from basic input conditions (triangle operators) proceeding through to system outputs. While, in general, the physical arrangement and interconnections of the system hardware in the system schematic can be parroted by the analyst in developing the GO chart, there are instances where the system logic and the system schematic arrangement are not compatible.

In Step 4, the analyst chooses the appropriate operator types to assign to the functional model. In this step, the logic that defines the specific success or failure state of interest in the model (as shown by the addition of the Type 11 in Figure 4-9(4)) is included. Steps 3 and 4 cannot be accomplished independently. The analyst must have a general idea in mind of the operator type that will be assigned to each element of the functional chart in order to draw it. However, there is some limited latitude in the assignment of operator types after completion of the functional chart.



1. DEFINE SYSTEM

2. ESTABLISH INPUTS/OUTPUTS

INPUTS	OUTPUTS
TANK	LCV A1
EP	LCV A2
ACT	LCV B1
	LCV B2

3. DRAW FUNCTIONAL GO CHART

4. DEFINE OPERATOR TYPES

5. DEFINE OPERATOR KINDS

6. DEFINE SIGNAL SEQUENCE

FIGURE 4-9. GO MODELING STEPS

In Step 5, the kind numbers are assigned to each operator requiring a probability assignment. Like components will generally be assigned identical kinds.

In Step 6, the analyst defines the sequencing to be used by the GO computer codes to quantify the model. The sequencing rule is simply that a given operator cannot be analyzed until the operators producing the inputs to the given operator have been analyzed. Within this constraint, there are several possible usable ordering arrangements.

4.4 CAUSE-CONSEQUENCE ANALYSIS

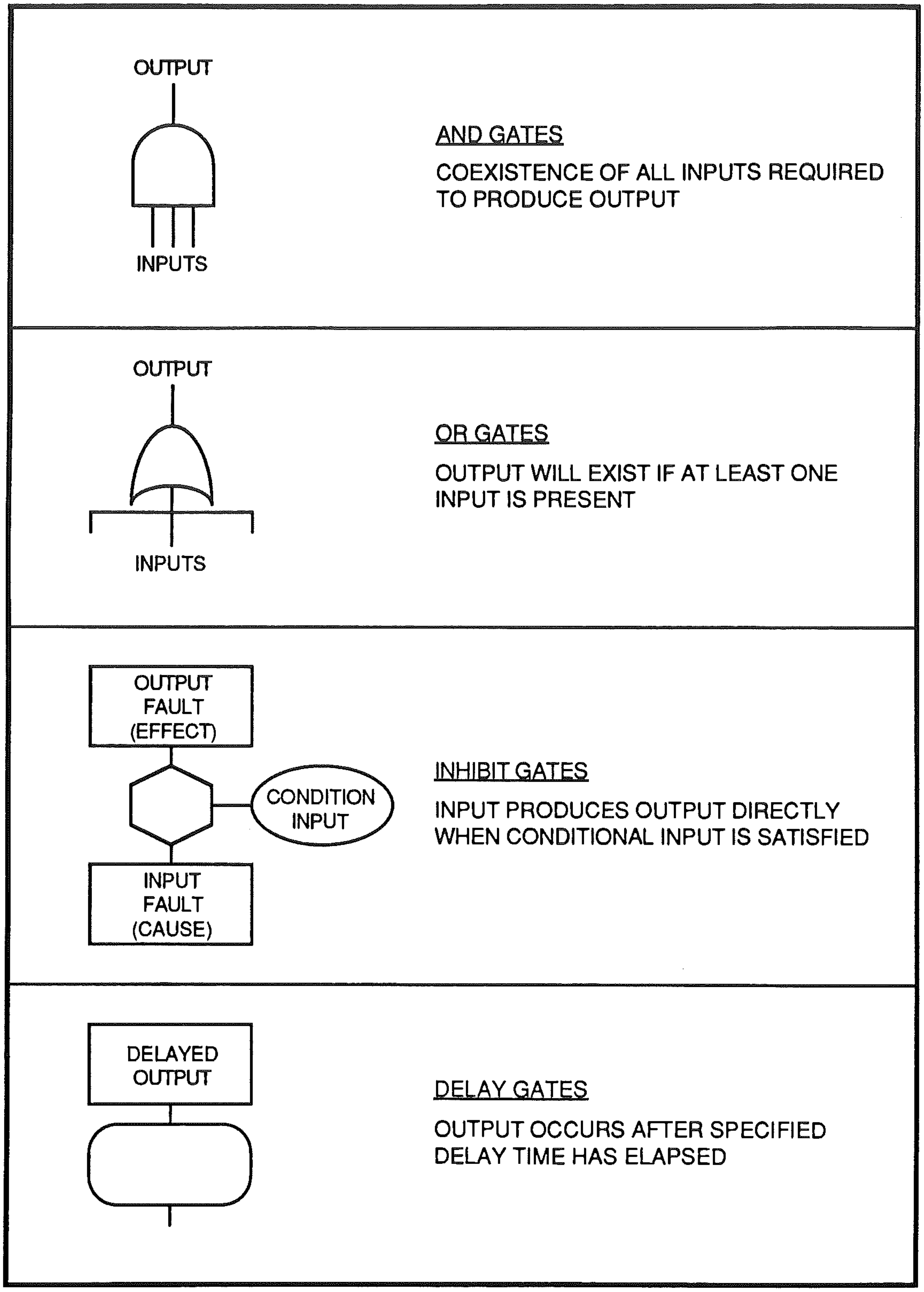
Cause-consequence analysis is a combination of fault tree analysis (cause) and consequence analysis.

The cause portions of the cause-consequence diagram are fault trees with their Top Events being component or system failures that can lead to various levels of undesired consequences, depending on the degree of mitigation provided by standby (backup, or safety) systems. The consequence portion of the diagram illustrates the array of possible consequence levels as a function of the binary state (failed or unfailed) of a standby system. The diagram, complete with cause and consequence portions, is referred to as the cause-consequence diagram. The well-constructed cause-consequence diagram provides a clear but often very detailed flow chart that illustrates system interrelationships that either preclude or contribute to the probabilities of occurrence of the various consequences possible to arise from a particular main Top Event called an "initiating event."

Symbols used in the cause portion of the cause-consequence diagram, given in Figures 4-10 and 4-11, are standard fault tree symbols. Symbols used in the consequence portion are given in Figure 4-12. The use of these symbols is demonstrated in a sample cause-consequence diagram in Figure 4-13.

4.4.1 Procedure for Constructing Cause-Consequence Diagrams

The procedure for constructing cause-consequence diagrams is described below. A more complete description is given by Nielsen (1975).



AND GATES

COEXISTENCE OF ALL INPUTS REQUIRED TO PRODUCE OUTPUT

OR GATES

OUTPUT WILL EXIST IF AT LEAST ONE INPUT IS PRESENT

INHIBIT GATES

INPUT PRODUCES OUTPUT DIRECTLY WHEN CONDITIONAL INPUT IS SATISFIED

DELAY GATES

OUTPUT OCCURS AFTER SPECIFIED DELAY TIME HAS ELAPSED

FIGURE 4-10. FAULT TREE LOGIC SYMBOLS

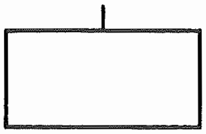
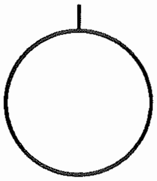
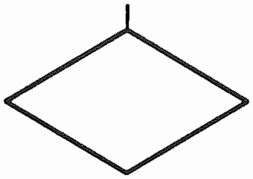
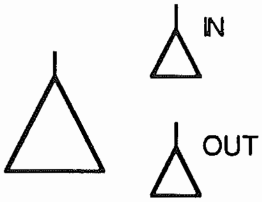
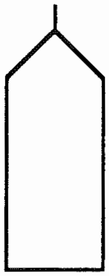
	RECTANGLE A FAULT EVENT RESULTING FROM THE COMBINATION OF MORE BASIC FAULTS ACTING THRUOUGH LOGIC GATES
	CIRCLE A BASIC COMPONENT FAULT - AN INDEPENDENT EVENT
	DIAMOND A FAULT EVENT NOT DEVELOPED TO ITS CAUSE
	TRIANGLE A CONNECTING EVENT OR TRANSFER SYMBOL
	HOUSE AN EVENT THAT IS NORMALLY EXPECTED TO OCCUR OR TO NEVER OCCUR. ALSO USEFUL AS A "TRIGGER EVENT" FOR LOGIC STRUCTURE CHANGE WITHIN THE FAULT TREE.

FIGURE 4-11. FAULT TREE LOGIC SYMBOLS (CONTINUED)

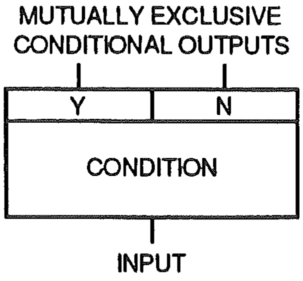
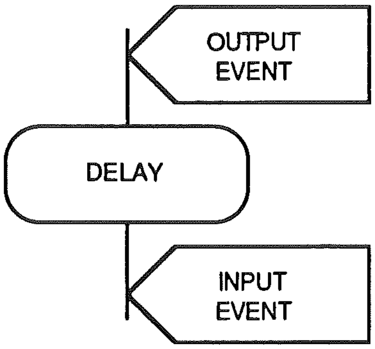
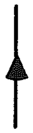
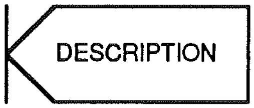
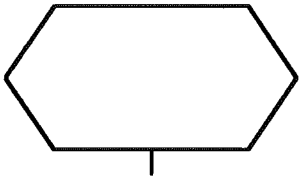
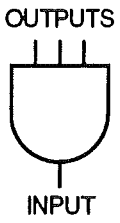
 MUTUALLY EXCLUSIVE CONDITIONAL OUTPUTS Y N CONDITION INPUT	BRANCHING OPERATOR OUTPUT IS "YES" IF CONDITION IS MET; :NO" OTHERWISE
	DELAY OPERATOR INDICATES THE AMOUNT OF TIME DELAY REQUIRED FOR OUTPUT EVENT TO RESULT FROM THE INPUT EVENT
	DIRECTOR INDICATES THE DIRECTION OF EVENT FLOW
	EVENT DESCRIPTOR DESCRIBES THE EVENT PRESENT AT SPECIFIED POSITION IN CHART
	CONSEQUENCE DESCRIPTOR DESCRIBES THE CONSEQUENCE. A TERMINAL SYMBOL
	INVERSE AND GATE ALL OUTPUTS OCCUR IF THE INPUT OCCURS

FIGURE 4-12. CAUSE-CONSEQUENCE DIAGRAM SYMBOLS

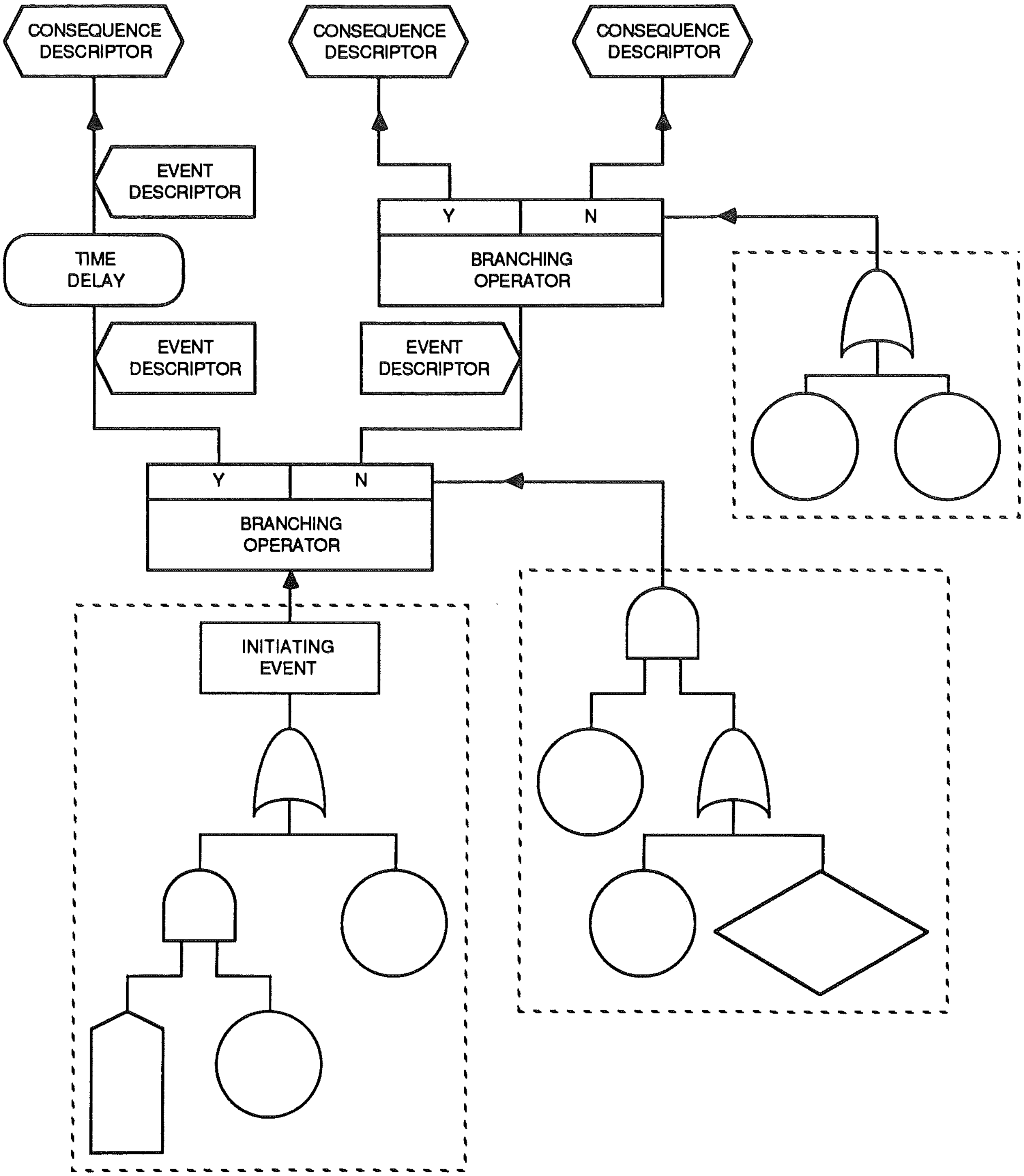


FIGURE 4-13. SAMPLE CAUSE-CONSEQUENCE DIAGRAM

The consequence diagram construction procedure begins with a choice of one or more "critical events." Critical events are changes in the system or facility state which provide a convenient starting point for the analysis. Examples of critical events are:

- A disturbance of an important system or facility variable, such as pressure, temperature, voltage, concentration, or speed.
- An event that leads directly to changes in several variables, such as opening or closing of a valve.
- An event that leads directly to activation of a safety system.
- A failure of some critical supply, such as electric power, cooling water, compressed air, or ventilation.
- A failure in control, or an error in manual operation or in maintenance.
- A breach of a pressure boundary, or of a boundary retaining toxic or inflammable material.

For each critical event that is identified, a consequence diagram is built up. Often, though, different critical events have similar consequences. In this case, two consequence diagrams may be merged.

The principal of consequence analysis is to trace the chains of events starting from the critical events. The chains of events are traced along pipes and wires, for example, of a facility, and along ventilation ducts and other paths by which a critical event can affect other parts of the facility. At each stage the changes resulting from the critical event are recorded on the consequence diagram.

Pipes, wires, ventilation ducts, and the spaces between components form the causal paths of the facility. Where there are several causal paths leading from a component, the event chains leading from the critical event may divide, giving rise to parallel event chains. Several sets of changes in facility state may take place side-by-side, independently of each other.

In some cases, the course which a chain of events takes will depend on the component state. For example, whether or not a fire alarm will activate an alarm bell will depend on whether or not the bell is working. Whether or not a pump will begin pumping water will depend on whether or not there is any water in the supply tank to be pumped. In such cases, alternative event chains will be possible and are added to the consequence diagram via a "condition box" for a branching operator (see Figures 4-12 and 4-13).

Chains of events are traced until either no further events of significance can be found, or some final consequence is reached. Final consequences are recorded on the diagram in "event boxes" providing consequence descriptors.

A different approach to consequence diagram construction is used for analyzing the possibilities of failures in operating procedures such as start up, shut down, batch production, or various measurement and maintenance procedures such as boiler blow down. The actions to be carried out are written down as a chain of events. Into this chain, the effects of the actions on the facility are interspersed, so that the result is a chain of events of the form

ACTION - EFFECT ON FACILITY - ACTION - EFFECT ON FACILITY - ACTION —

This chain is then modified by considering what happens if the required actions are not carried out, and what the effects of the actions are if the facility is not in its normal state. The event boxes are made into decision boxes with a "yes" and a "no" exit. The consequences of omitting an action, of performing a wrong action, or of performing a correct action when there is a latent (undetected) failure in the facility, are then evaluated using the consequence analysis procedure given above.

As was shown conceptually in Figure 4-13, fault tree procedures are then needed to investigate the causes of critical events. Fault trees can also be added to the consequence chain decision boxes to explain the reasons why some components do not respond properly, for example, as a result of latent failures. In this way, a full cause-consequence diagram can be built up.

Some special features of cause-consequence diagrams are:

- Delays can be placed in the event sequences.
- Initiating events with identical event sequences can be joined together by OR gates (as in Figure 4-14).

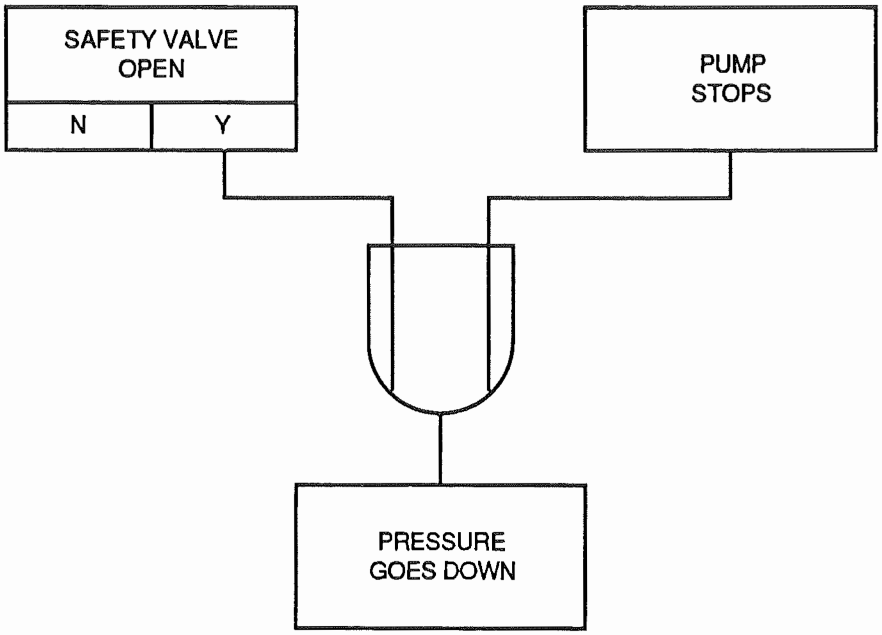


FIGURE 4-14. OR GATE

- A facility state that is not a final consequence state, such as"waiting for repair," can be represented as a variable delay with a triangle symbol (see Figure 4-15). A small black triangle is used to show different ways in which the waiting state can end.

4.4.2 A Tutorial Example

The following example taken from Fussell et al. (1976) is presented to demonstrate some of the fundamental aspects of cause-consequence analysis. The sample system is shown in Figure 4-16. The motor is located such that it has a chance of causing a catastrophic fire. Figure 4-17 is the cause-consequence diagram for this example situation.

The initiating or trigger event is "motor overheats." This is the Top Event of the fault tree at the bottom of Figure 4-17. This fault tree is the primary cause portion of the cause-consequence diagram.

The consequence portion of the cause-consequence diagram unfolds to reflect the sequence of events that could be encountered by the system, beginning with the initiating event. It uses branching operators, time delays, OR gates, and inverse AND gates, with event descriptor tags appropriately placed to add clarity. This process continues until each path ends in a consequence description. Each branching operator has its associated fault tree that develops causes for the undesirable condition in that operator. Common cause failure exposure potential originates among these fault trees for branching operators and the initiating event.

4.5 DIRECTED GRAPHS

The digraph-fault tree methodology was devised by Lapp and Powers (1977a and 1977b) for safety analyses of chemical processing systems. Further development of the technique is described by Allen and Rao (1980). The directed graph, or digraph, is a multivalued logic diagram that describes the interrelationships among process variables. In addition, events that change or nullify the relationships among process variables are shown. These events appear as basic events in the fault tree developed from the digraph. In essence, the digraph is an intermediate step between the system schematic and the construction of a fault tree.

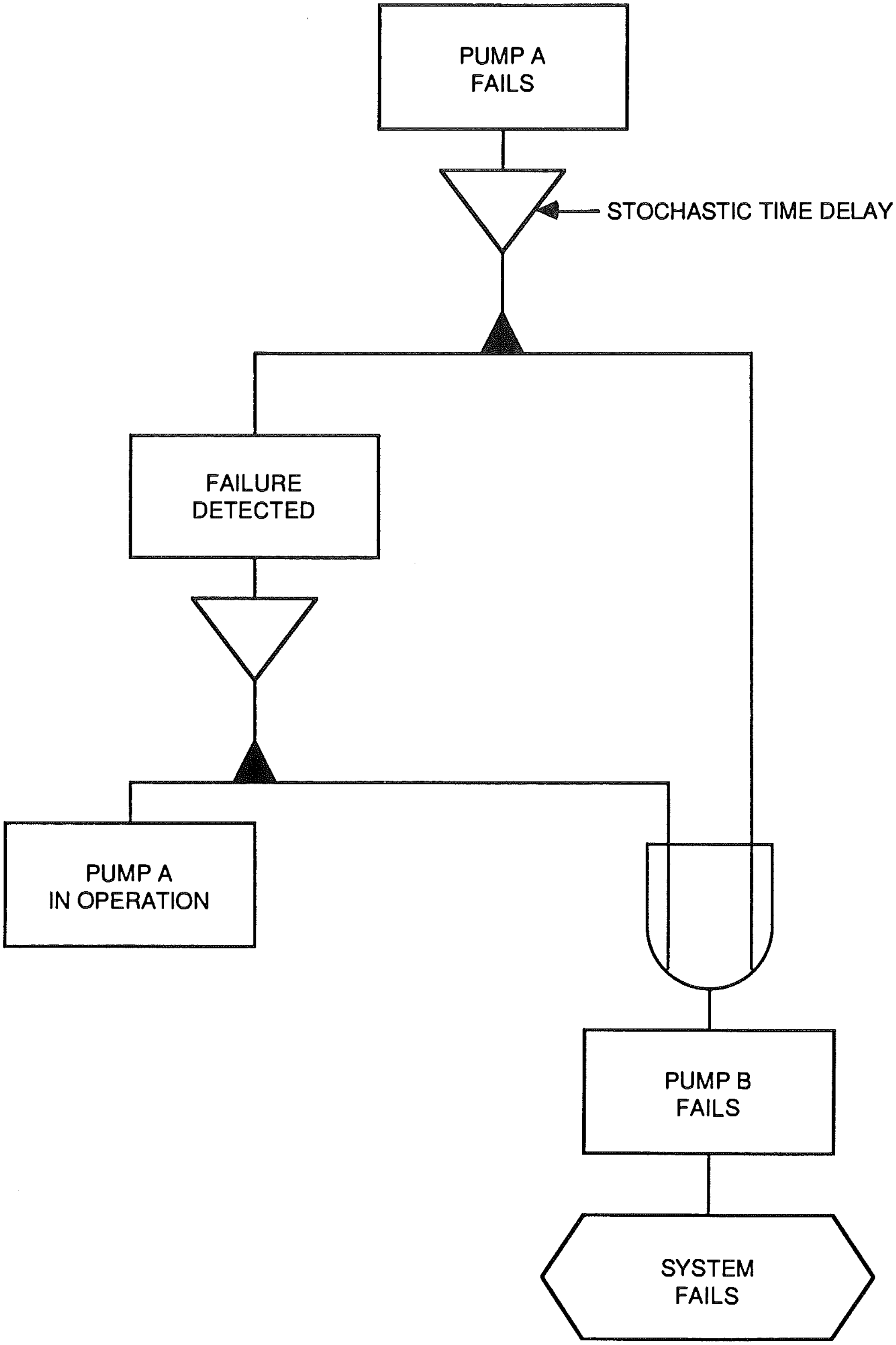
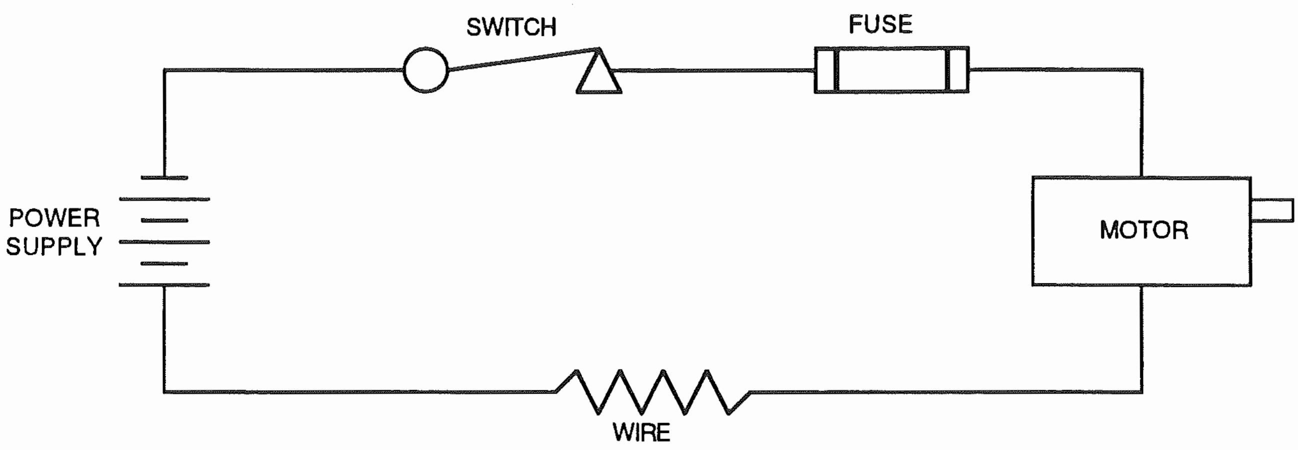


FIGURE 4-15. CONSEQUENCE DIAGRAM WITH TIME DELAY



TOP EVENT	=	MOTOR OVERHEATS
INITIAL CONDITION	=	SWITCH CLOSED
NOT-ALLOWED EVENTS	=	FAILURES DUE TO EFFECTS EXTERNAL TO SYSTEM
EXISTING EVENTS	=	SWITCH CLOSED

FIGURE 4-16. SAMPLE SYSTEM

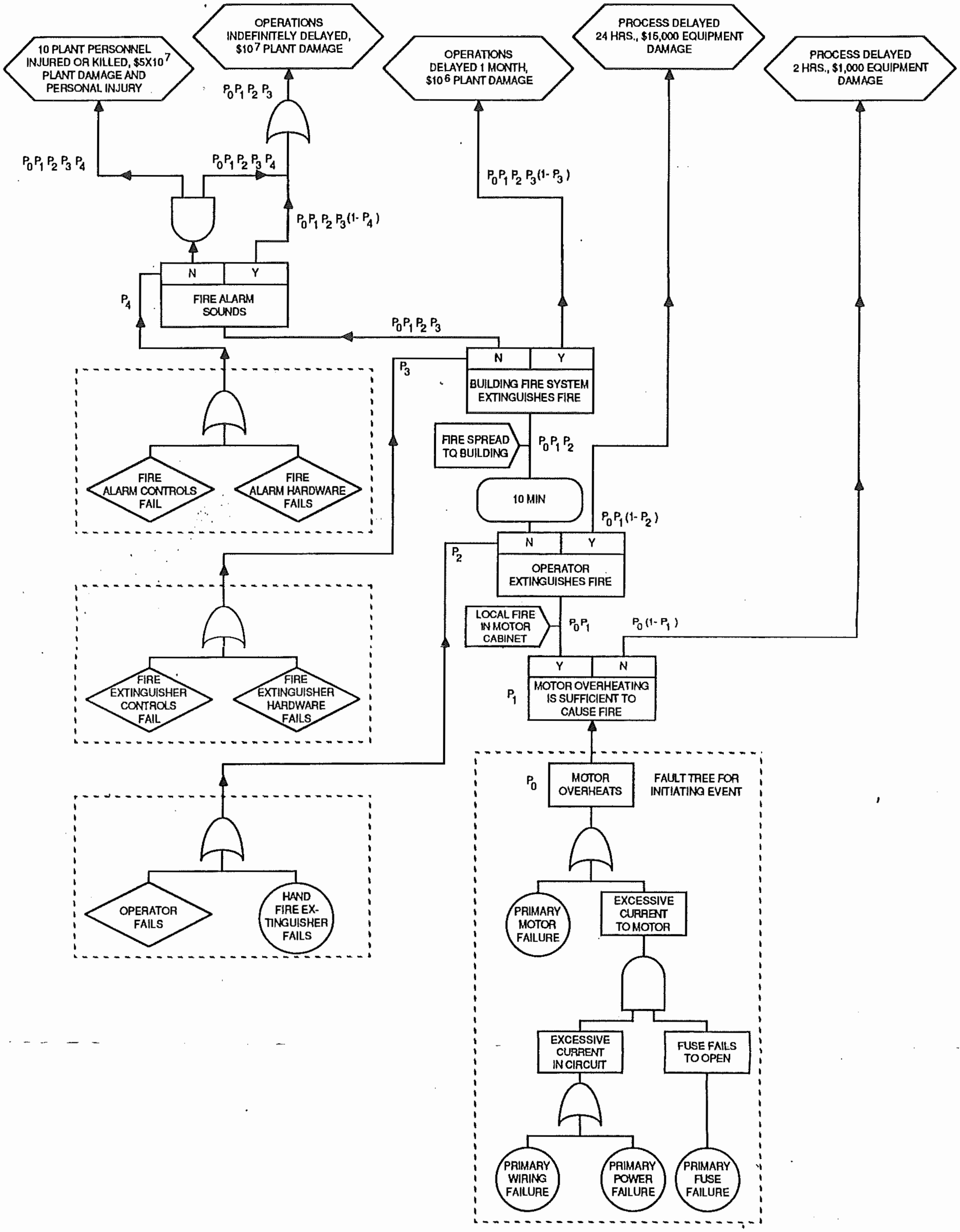

10 PLANT PERSONNEL INJURED OR KILLED, $5X10^7$ PLANT DAMAGE AND PERSONAL INJURY
OPERATIONS INDEFINITELY DELAYED, 10^7 PLANT DAMAGE
OPERATIONS DELAYED 1 MONTH, 10^6 PLANT DAMAGE
PROCESS DELAYED 24 HRS., $15,000 EQUIPMENT DAMAGE
PROCESS DELAYED 2 HRS., $1,000 EQUIPMENT DAMAGE
$P_0P_1P_2P_3$
$P_0P_1P_2P_3P_4$
$P_0P_1P_2P_3P_4$
$P_0P_1P_2P_3(1-P_3)$
$P_0P_1P_2P_3(1-P_4)$
N
Y
FIRE ALARM SOUNDS
P_4
$P_0P_1P_2P_3$
P_3
N
Y
BUILDING FIRE SYSTEM EXTINGUISHES FIRE
FIRE ALARM CONTROLS FAIL
FIRE ALARM HARDWARE FAILS
FIRE SPREAD TO BUILDING
$P_0P_1P_2$
10 MIN
$P_0P_1(1-P_2)$
P_2
N
Y
OPERATOR EXTINGUISHES FIRE
FIRE EXTINGUISHER CONTROLS FAIL
FIRE EXTINGUISHER HARDWARE FAILS
LOCAL FIRE IN MOTOR CABINET
P_0P_1
$P_0(1-P_1)$
Y
N
P_1
MOTOR OVERHEATING IS SUFFICIENT TO CAUSE FIRE
P_0
MOTOR OVERHEATS
FAULT TREE FOR INITIATING EVENT
OPERATOR FAILS
HAND FIRE EXTINGUISHER FAILS
PRIMARY MOTOR FAILURE
EXCESSIVE CURRENT TO MOTOR
EXCESSIVE CURRENT IN CIRCUIT
FUSE FAILS TO OPEN
PRIMARY WIRING FAILURE
PRIMARY POWER FAILURE
PRIMARY FUSE FAILURE

FIGURE 4-17. SAMPLE SYSTEM CAUSE-CONSEQUENCE DIAGRAM

Ideally, in this procedure, one starts with the basic laws of mass, energy, and momentum and constructs differential and algebraic equations that describe relationships that exist among system variables, e.g., temperature, mass flow, pressure. The digraph procedure takes a continuum of possible values of the variables and uses discrete logic to model the continuum with functional models. These models are useful in failure analysis. For example, complex laws describe heat balances for heat exchangers. A functional model indicates that a decrease of the flow of cooling water to a heat exchanger will cause an increase in the output temperature of the hot stream. These functional relationships are embodied in digraphs.

The digraph procedure was devised primarily for failure analysis of control systems. Manual fault tree techniques in general do not work well in modeling control systems because it is difficult to envision the topology of the system control loop structure from the system schematic when manually constructing a fault tree. This structure may exclude events that can occur in the fault tree at the "local" level — local in the sense of what is immediately necessary and sufficient to cause an event. These logic consistency checks are an important part of the procedure.

The digraph clearly displays the system control loops. Its structure describes how variables are linked at the system level. The synthesis algorithm, that transforms the digraph into a fault tree for a specified Top Event, requires that the control loop structure in the digraph be found. Finding the loop structure allows two important steps to be conducted in the digraph-fault tree procedure:

- By knowing the dynamics of the relationships between the variables, one can assess the dynamics of the response of the control loops.
- One can locate trigger nodes on the control loops. These nodes define the operators to be used. These operators are used as"templates" that transform the digraph into the fault tree.

The digraph-fault tree procedure consists of two basic steps that are particularly efficient in the analysis of complex systems.

> *Step One* - Identification of (1) the cause and effect type of relationship between variables (edges in the diagram) and (2) "unusual" system states (component failures, basic event states, etc.). This information is displayed in the system digraph.

Step Two - Construction of the fault tree via a digraph-to-fault tree synthesis algorithm. The synthesis algorithm accounts for the control loop structure in its entirety when constructing the fault tree.

Conventional fault tree synthesis requires the analyst to consider all aspects simultaneously.

The digraph-fault tree procedure also gives a systematic format for the analyst to follow in considering the failure modes and unusual states for all components and variables.

Another useful feature of the digraph-fault tree procedure is that multivalued logic can be considered. Variables are "discretized" into five possible values, namely: normal, moderate (high or low), and large (high or low). Values of variables other than normal values are called disturbances; they are similar to "perturbation variables" in control theory. Control loops are classified according to their ability to cancel a disturbance of a given size or vice versa. Control loops may fail to cancel disturbances because control devices are inactive (e.g., control devices fail in the "stuck" mode) or the control loops may be the cause of the disturbance if control devices fail high or low. Hence, classifying failure modes of control devices is also an important part of the procedure.

As with fault trees, digraphs are always constructed with a Top Event in mind.

The digraph for pressure tank rupture is shown in Figure 4-18. The variables P_{Tank} and I denote tank pressure and current, respectively. There are two negative feedback loops (NFBLs) indicated by bold cyclic lines in the digraph. The function of these loops is to counteract the effect of the disturbance, "timer contacts fail to open," which causes the pump motor to continue operating, resulting in overpressure. There are two NFBLs:

- The operator sensing indicated pressure and opening the manual switch if pressure is too high (similar in scope to an interlock or trip function) (NFBL #1).
- Pressure relief valve opening in the event of excess pressure (NFBL #2).

According to the digraph terminology, events that inactivate the loops are called zero gain events and appear as basic events in the fault tree. The value of the gain appears on an arrow connecting the nodes that are system variables. Gain has the meaning Y/ X where Y is the dependent variable and X the independent variable. A gain of +1 or -1 implies a moderate positive or negative relationship between variables, ±10 indicate a strong relationship, and a gain of zero implies no relationship.

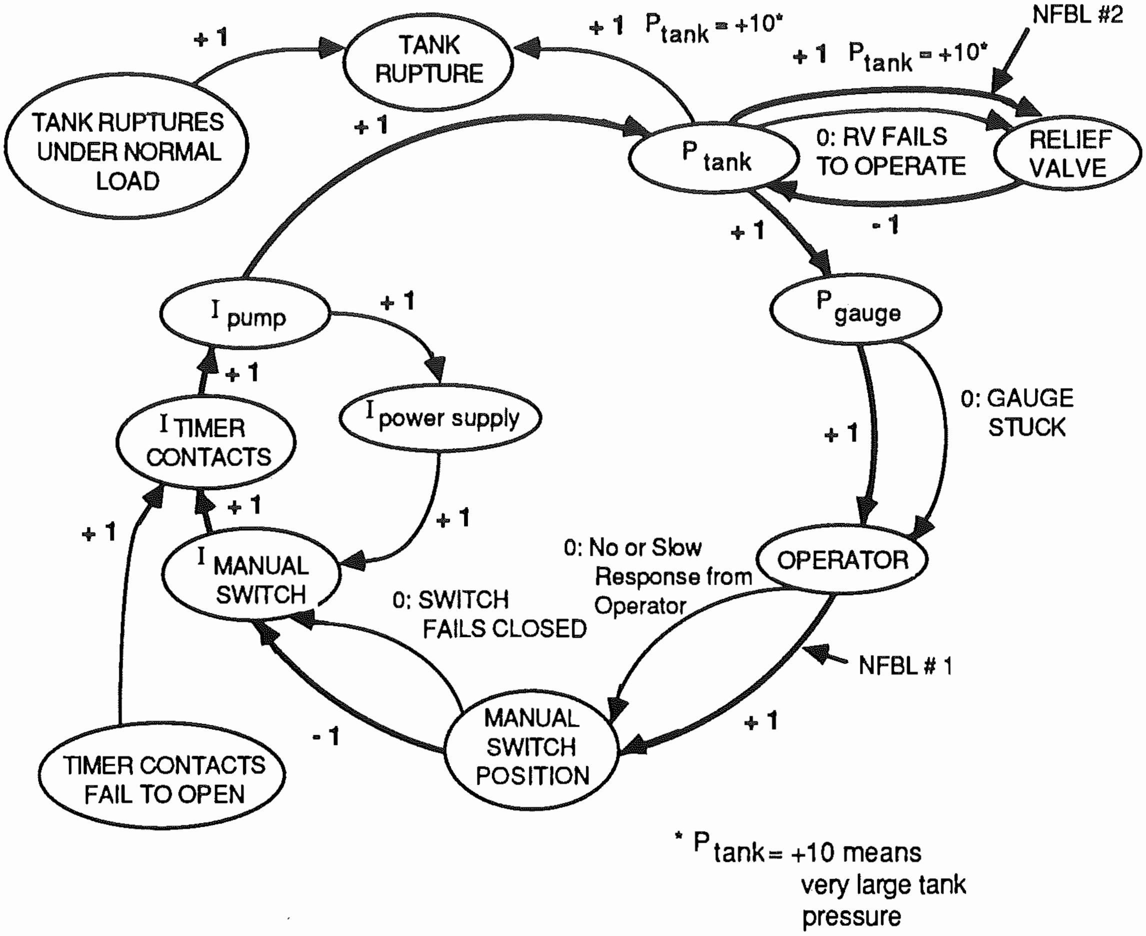


FIGURE 4-18. DIGRAPH FOR PRESSURE TANK RUPTURE

The fault tree is generated from the digraph via the synthesis algorithm that delineates how a control loop can either cause or pass a disturbance, resulting in the occurrence of the Top Event. The resulting fault tree is shown in Figure 4-19. There is a total of six basic events (basic events represent the limit of resolution in fault tree analysis) — two of which are initiating events:

- Tank ruptures under normal load
- Timer contacts fail to open

The tank rupturing under normal load is an initiating event and there are no mitigative features in the system to prevent this event from occurring. The timer contacts failing to open is another initiating event since it can initiate an event sequence leading to overpressure. The remaining four basic events,

- Relief valve stuck (fails pressure protection)
- Pressure gauge stuck (fails operator shutdown)
- No or slow response from operator (fails operator shutdown)
- Manual switch fails to open (fails operator shutdown),

do not cause overpressure but do inactivate the NFBLs.

Complicated control systems may contain nested loops. Such loops would occur with cascade control. Constructing a digraph, as was shown in Figure 4-18, helps the analyst understand how system variables are linked together at the system level and how disturbances may propagate through the system to cause a defined Top Event to occur.

The thrust vector control system for the Space Shuttle is a complex system that is important in preventing loss of either the vehicle or crew, particularly with loss of one main engine. Also, loading of liquid hydrogen and oxygen during ground operations is a highly automated process. The digraph would be a useful tool in assessing control system failures for these functions.

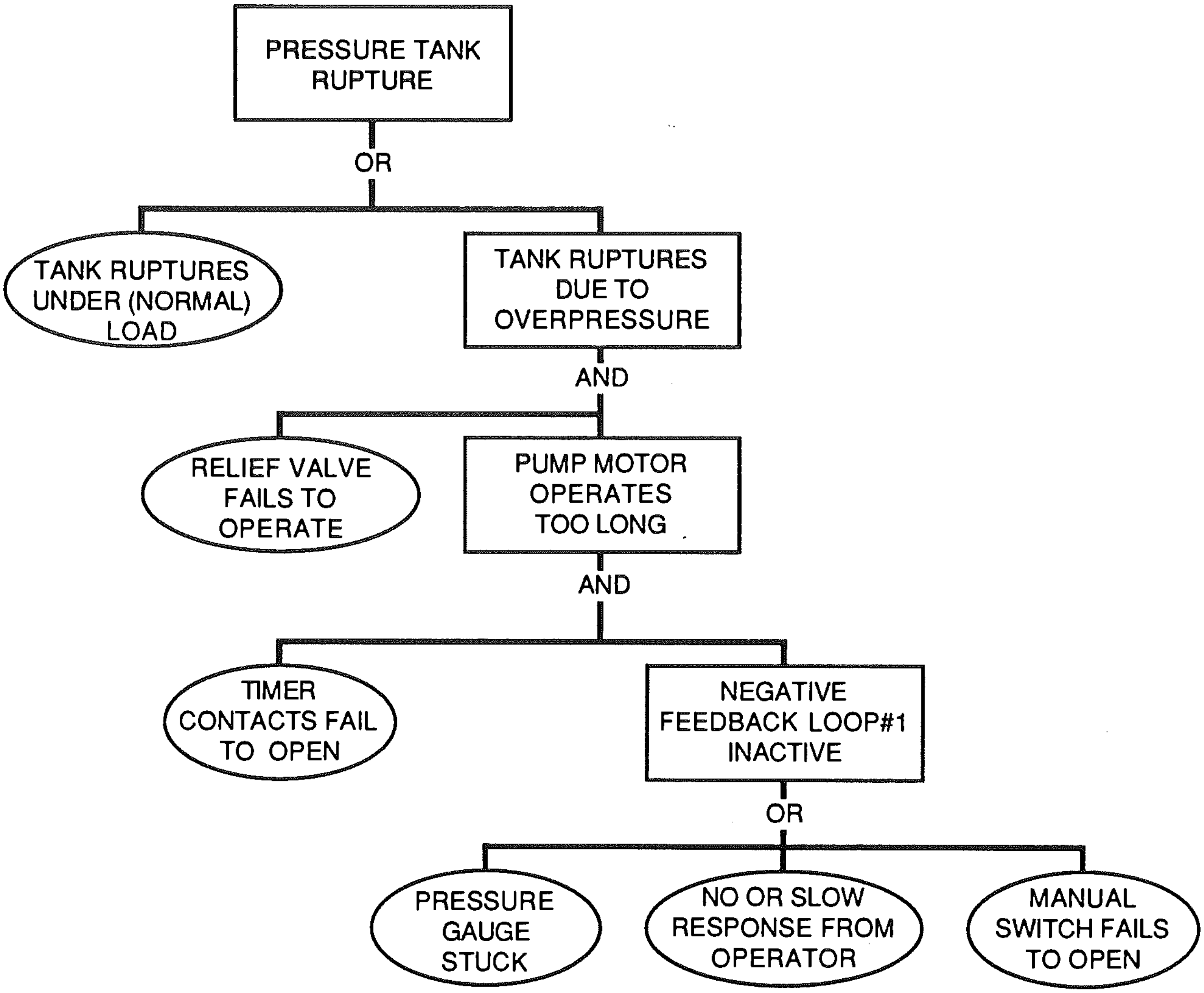

PRESSURE TANK RUPTURE
OR
TANK RUPTURES UNDER (NORMAL) LOAD
TANK RUPTURES DUE TO OVERPRESSURE
AND
RELIEF VALVE FAILS TO OPERATE
PUMP MOTOR OPERATES TOO LONG
AND
TIMER CONTACTS FAIL TO OPEN
NEGATIVE FEEDBACK LOOP#1 INACTIVE
OR
PRESSURE GAUGE STUCK
NO OR SLOW RESPONSE FROM OPERATOR
MANUAL SWITCH FAILS TO OPEN

FIGURE 4-19. FAULT TREE FOR PRESSURE TANK RUPTURE

4.6 COMPARATIVE STRENGTHS AND WEAKNESSES OF THE SEVERAL LOGIC MODELS

A traditional approach used within NASA is failure modes and effects analysis (FMEA). The advantage of FMEA is that it is easy to apply. However, it is inefficient in identifying and assessing multiple component failures. The risk models that have been described have the capability of considering multiple failures. The strengths and weaknesses of each model are next addressed.

4.6.1 Conventional Fault Tree Analysis

Conventional fault tree analysis (FTA) has been in existence since the early 1960's. It is well understood and several textbooks and many articles have been written on fault tree metholodogy and its applications. The nature of FTA requires an interrogatory thought process. The analyst must understand how the system works and how the system fails in a manner such as to contribute to the occurrence of the Top Event. One of the advantages of FTA is the qualitative insight into system behavior that is gained when the fault tree is constructed.

A disadvantage of fault tree analysis and its development process is that different analysts may construct different fault trees for the same application. However, by conducting an importance analysis to find the dominant mincutsets, the differences between the fault trees can be reconciled. Another disadvantage is that the fault tree does not resemble the system schematic. In many cases, fault trees are not scrutable. To alleviate this problem, an identification scheme that cross-references the event description to the system schematic is helpful. Finally, fault tree analysis can be a very time-consuming process. However, FMEA can also be time-consuming, particularly when numerous unimportant failure modes are considered. Fault tree analysis considers only those failure modes that contribute to a defined undesired Top Event.

4.6.2 Event Trees

Event trees are conceptually simple. The analysis starts with a defined initiating event and considers the subsequent combination of successes and failures of various systems and system functions that become involved. Event trees are useful for displaying accident scenarios (1) with complex time dependencies and (2) with complex interrelationships of system functions. The event tree headings refer to success or failure of a particular system function. The analyst must decide on the system ordering in the event tree headings. In theory, with N systems, there are 2^N possible orderings. However, there are many illogical combinations that can be eliminated when the analyst considers timing and dependency of events. Event trees can become large and difficult to follow.

4.6.3 GO Methodology

As described in the PRA Procedures Guide (Nuclear Regulatory Commission, 1983), some key features of the GO method are the following: (1) models follow the normal process flow; (2) model elements have almost one-to-one correspondence with system elements and handle most component and system interactions and dependencies; (3) models are compact and easy to validate; (4) outputs represent both success and failure states; (5) models can be easily altered and updated; (6) fault sets can be generated without altering the basic model; (7) system operational aspects can be incorporated; and (8) numerical errors due to truncation are known and controlled.

However, in some cases the GO chart may not bare its relationship to the system schematic so clearly. For example, take the example of a pressure tank system with two timers. These timers are in series with regard to reliability (i.e., starting the pump motor) and parallel with regard to safety (i.e., turning the pump motor off). Employing success (failure) logic, an AND (OR) and (AND) OR gate are generated, respectively. Success logic is used by GO, failure logic is used in fault tree analysis. The important point to be made is that the logic model depends upon the system event being modeled and the logic can differ from the system schematic. Another example is flow diversion through a path that contains two valves in series. When the analyst constructs the GO chart, the desired event of interest is flow through the regular path and no flow through the diversion path — the latter event is represented by OR logic instead of AND logic. Another disadvantage to a GO chart is that it can be inscrutable for a complex system, possibly as much as a fault tree.

The GO methodology is effective in predicting system reliability and availability for systems with well-defined sets of components and events able to be represented in a system schematic. It is not good for modeling abstract failures such as software logic errors and secondary failures such as fire and explosion, or out-of-tolerance failures such as are caused by excessive environmental or operational stress placed upon a component. Fault tree analysis models these failures straightforwardly. Also, GO is not oriented to evaluating specific Top Events' occurrence frequencies or probabilities, the quantitative figures of merit it is generally desired to evaluate, whereas fault tree analysis does produce these occurrence frequencies or probabilities, computed on a per unit time or a per demand basis. Furthermore, fault duration times of events such as latent failures or pre-existing conditions which may change if a new Top Event is chosen, are important and can be modeled with fault trees. The GO methodology in its current form is not suited for calculations in which fault duration times of enabling events must be specified.

4.6.4 Cause-Consequence Diagrams

Cause-consequence analysis is useful for especially detailed analysis of complex activities. It can consider in one diagram parallel sequences that cannot be considered in a single event tree; an event tree must be constructed for each sequence. Event descriptions concerning timing, conditionality of events, and sequential effects are easy to follow in cause-consequence diagrams. On the other hand, the logic in event trees may be difficult to follow for complex time sequences. Often, event trees must be accompanied with separate complex descriptions.

The disadvantage of cause-consequence diagrams is that they can become too large very quickly. Also (in common with digraphs), they suffer from the fundamental disadvantage that they are not well understood by the risk analysis community at large.

4.6.5 Directed Graphs

Digraphs may be used to supplement fault trees in the analysis of control systems. An advantage of digraphs is that they treat multivalued logic and timing, and display the topology of system variables. Digraphs are particularly useful when the analyst keys in on a system variable that is sensed or manipulated by one or more control loops. Digraphs are good for studying the causes of upset conditions that can be initiating events in the analysis.

Disadvantages to digraphs are that they are time consuming to generate and they are not well understood by the risk analysis community. In addition, fault tree algorithms must be devised for various control loop configurations. The number of algorithms can be large, reflecting the fact that to model control system faults requires a more sophisticated thought process than conventional fault tree analysis.

4.7 QUALITATIVE ANALYSIS

The next step after constructing a logic model is qualitative analysis. It entails finding the minimal cut sets (mincutsets) and performing other Boolean algebraic operations. Mincutsets are sets of basic events whose occurrence ensures the occurrence of the Top Event.

4.7.1 Mincutsets

As shown in Table 4-2, there is a total of five mincutsets in the pressure tank rupture fault tree that was given in Figure 4-2. These mincutsets also describe how accident sequence PO*OS*PP can occur. Each mincutset contains one initiating event, which implies there is only one time sequence by which a mincutset can cause system failure. Three initiating events are listed:

- tank ruptures under normal load
- voltage surge (a common-cause initiating event, see below)
- timer contacts fail to open.

4.7.2 Common Cause Analysis

Common cause analysis deals with identifying and evaluating operational or environmental conditions that can simultaneously fail two or more otherwise independent system components. Such conditions include, for example, impact, flood, fire, or common maintenance faults, and are called secondary failures in fault tree analysis. An example of a common cause failure for the Space Shuttle is cavitation of a high pressure oxidizer turbopump which can be caused by an abrupt cutoff of liquid oxygen. Cavitation can cause missile generation that can fail the other main engines of the Space Shuttle and lead to the loss of the vehicle and crew.

A method for conducting common cause analysis is to construct fault trees that exclude common cause failures and generate their mincutsets. The mincutsets are then searched to determine if they have a common cause susceptibility. The advantage to this approach is that a smaller fault tree is generated. For example, voltage surge, that was shown as a basic event in the fault tree in Figure 4-2, is an example of a secondary failure. If this event had been excluded from the fault tree, then mincutset number 2 in Table 4-2 would not be generated but would be generated later when mincutset number 5 is searched for a common-cause dependency. In this case, it is recognized that a voltage surge would simultaneously weld shut the timer contacts and the manual switch contacts.

Mincutset	Description
1	• Tank Ruptures Under Normal Load (i)
2	• Voltage Surge (i) • Relief Valve Fails to Operate (e)
3	• Timer Contacts Fail to Open (i) • Relief Valve Fails to Operate (e) • Pressure Gauge Stuck (e)
4	• Timer Contacts Fail to Open (i) • Relief Valve Fails to Operate (e) • No or Slow Operator Response (e)
5	• Timer Contacts Fail to Open (i) • Relief Valve Fails to Operate (e) • Manual Switch Fails to Open (e)
Notes:	(i) Denotes an initiating event (e) Denotes an enabling event

Table 4-2. Listing of Mincutsets
(for Fault Tree in Figure 4-2)

4.7.3 Boolean Factorization

In Table 4-3, the mincutsets are factored in terms of basic events. The terms in brackets define the critical system states for the occurrence of the initiating event. Stated qualitatively, critical system states describe the vulnerability of the system to the occurrence of the initiating event.

4.7.4 Structural Importance

A qualitative measure of importance is called structural importance (Birnbaum, 1969). The following series-parallel system is used to demonstrate how to compute structural importance.

Expression
Number

1 Pressure Tank Rupture =

$$\text{Tank} + \left[\left\{ \text{R - Valve} \right\} * \text{V - Surge} \right] + \left[\left\{ \text{R - Valve} * \left(\text{Gauge} + \text{Operator} + \text{Switch} \right) \right\} * \text{Timer} \right] \tag{4-1}$$

2 Second Top Event = . . . etc.

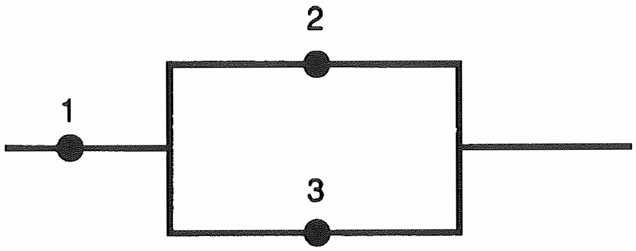
where

+	=	Boolean Union (OR)
*	=	Boolean Intersection (AND)

NOTES:

Expression 1 is the first Top Event's Boolean expression factored according to basic events. Boolean terms in braces define the critical system state for each initiating event.

TABLE 4-3. BOOLEAN FACTORIZATION OF TOP EVENTS' MINCUTSETS IN TABLE 4-2



The mincutsets for this system are (1) and (2, 3) where (1) denotes the failure of component 1, etc. It will be seen that component number 1 is individually qualitatively more important than components 2 or 4. If the state of one component is fixed, the remaining number of possible system states is 2^{n-1}, where *n* is the number of components. For this example, *n*=3 and the number of system states to be considered is $2^{3-1} = 2^2 = 4$. If, in particular, the state of component 1 is fixed, the remaining system states are:

SYSTEM STATE	2	3	CRITICAL SYSTEM STATE WHEN 1 FAILS?
1	WORK	WORK	YES
2	FAIL	WORK	YES
3	WORK	FAIL	YES
4	FAIL	FAIL	NO

Structural importance is defined as the fractional number of system states that are critical for a component. A critical system state for a component is defined as a system state such that the system makes a transition from the unfailed to the failed state when the component fails. The structural importance of component 1 is then 3/4 = 0.75. System state 4 is not a critical system state for component 1 since the system is already failed. For components 2 and 3, there is only one system state that is critical. Hence, the structural importance for each of components 2 and 3 is 1/4 = 0.25.

Examining Table 4-2, it is seen qualitatively that the tank rupturing under load is the most important event since it is a single event leading to tank rupture. However, note that this event is a passive failure, with necessarily low probability. Thus, from a quantitative or probabilistic viewpoint, it is not the most important failure. This implies that probabilistic importances can differ greatly from structural importances.

4.8 THE NECESSITY OF DISTINGUISHING INITIATING AND ENABLING EVENTS

Initiating events create a perturbation in a system variable that causes the Top Event to occur. Enabling events permit initiating events to occur; e.g., they represent failures of safety mitigation features. The distinguishing of initiating and enabling events is important from both a qualitative and probabilistic viewpoint. See Dunglinson and Lambert (1983) for a discussion of initiating and enabling events and their use in FTA.

Enabling events are of two types:

- pre-existing conditions such as
  - loss of system redundancy
  - conditions for fire and explosion
- demand failures.

Examples of each type of enabling event are given below.

4.8.1 Enabling Events for Preexisting Conditions

As was described in Section 4.1, for fire or explosion to occur all of the following conditions are needed (assuming reactants are below the autoignition temperature):

- heat or ignition source present
- flammable species present between the lower and upper flammability limits
- oxygen present above minimum concentration for combustion.

These conditions can occur in any time sequence. It is important to note that the event that occurs last is the initiating event since it causes the fire to occur. The other two events are enabling events that permit the fire to occur, given the occurrence of the last event.

The fault tree for fire and explosion for the pressure tank system was shown in Figure 4-4. There are seven mincutsets, as shown in Table 4-4.

All the basic events in Table 4-4 can also be enabling events and fault duration times must be assigned to these events.

MINCUTSET NUMBER	MINCUTSET DESCRIPTION			
1	LOSS OF NITROGEN PURGE (i)	IGNITION SOURCE PRESENT (i)	PRESSURE TANK RUPTURE UNDER LOAD (i)	
2	LOSS OF NITROGEN PURGE (i)	IGNITION SOURCE PRESENT (i)	LEAK IN PRESSURE TANK SYSTEM (i)	
3	LOSS OF NITROGEN PURGE (i)	IGNITION SOURCE PRESENT (i)	PRESSURE RELIEF VALVE STUCK OPEN (i)	
4	LOSS OF NITROGEN PURGE (i)	IGNITION SOURCE PRESENT (i)	VOLTAGE SURGE (i)	
5	LOSS OF NITROGEN PURGE (i)	IGNITION SOURCE PRESENT (i)	TIMER CONTACTS FAIL TO OPEN (i)	SWITCH CONTACT FAIL TO OPEN (e)
6	LOSS OF NITROGEN PURGE (i)	IGNITION SOURCE PRESENT (i)	TIMER CONTACTS FAIL TO OPEN (i)	NO OR SLOW OPERATOR RESPONSE (e)
7	LOSS OF NITROGEN PURGE (i)	IGNITION SOURCE PRESENT (i)	TIMER CONTACTS FAIL TO OPEN (i)	PRESSURE GAUGE READS LOW OR STUCK (e)

Notes: (i) Denotes an initiating event
(e) Denotes an enabling event

TABLE 4-4. MINCUTSETS FOR FIRE AND EXPLOSION

Assuming perfect inspection, the maximum fault duration time for the nitrogen purge system is 24 hours. The same can be said about the basic event "leak in pressure tank system." It is assumed, in this case, that the operator would detect flammable gas leakage when inspection occurs.

However, basic events that cause gross leakage from the system have a small fault duration time since the operator will notice that the tank pressure does not increase and he will shut the system down. The basic events that cause gross leakage are:

- pressure tank rupture under normal load
- pressure relief valve stuck open
- voltage surge
- timer contacts fail to open.

What the above statement implies is that fire or explosion is more likely to occur when the nitrogen purge system fails before the gross leakage event. Conversely, given a gross leakage, there is a small time window for the nitrogen purge system to fail before system shutdown. However, it may be said for the small leakage event, fire or explosion can occur in any time sequence.

This example provides other insights. Fault trees are always generated and analyzed with a Top Event in mind. Fault duration times are important in the probabilistic evaluation of fault trees. The distinction between initiating and enabling events is a necessary consideration.

Another example of a preexisting condition results from placing a second timer in series with the timer that was shown in Figure 3-1. In this case, redundancy has been incorporated to prevent the occurrence of overrun of the pump when a random failure of one timer occurs. One timer must fail first to create a condition necessary for the occurrence of a second event, the failure of the second timer (the initiating event), to cause system failure to occur.

4.8.2 Enabling Events for Demand Failures

For the second type of enabling event, demand failures can occur before, during, or after the occurrence of the initiating event. For example, failure of the relief valve occurs before or at the time of the demand to permit overpressure, whereas failure of the operator occurs after the demand.

The important point to be made is that enabling events have fault duration times associated with them, which means if their occurrence is detected before an associated demand is made, accidents can be prevented or mitigated.

4.8.3 Space Shuttle Examples

Figure 4-20 is a fault tree for high-pressure oxidizer turbopump (HPOTP) explosion or burnthrough. This event results in hot gaseous oxygen entering the aft fuselage area with the possibility of a pump or engine fire.

The reader is referred to the inputs below the top level OR gate in Figure 4-20, specifically, the second input "Intermediate purge seal failure causes pump fire." A failure of the purge seal system (left hand input of the AND gate) is the initiating event that in turn can be caused by any of the seven inputs listed below:

- helium supply fails
- helium line ruptures
- helium passage blocked in pneumatic control assembly
- filter blocked by contamination or ice
- purge solenoid valve armature or pushrod jammed
- both coils or harnesses fail in intermediate purge solenoid
- controller fails to energize or keep purge control valve open

Each of these seven events is an <u>initiating</u> event.

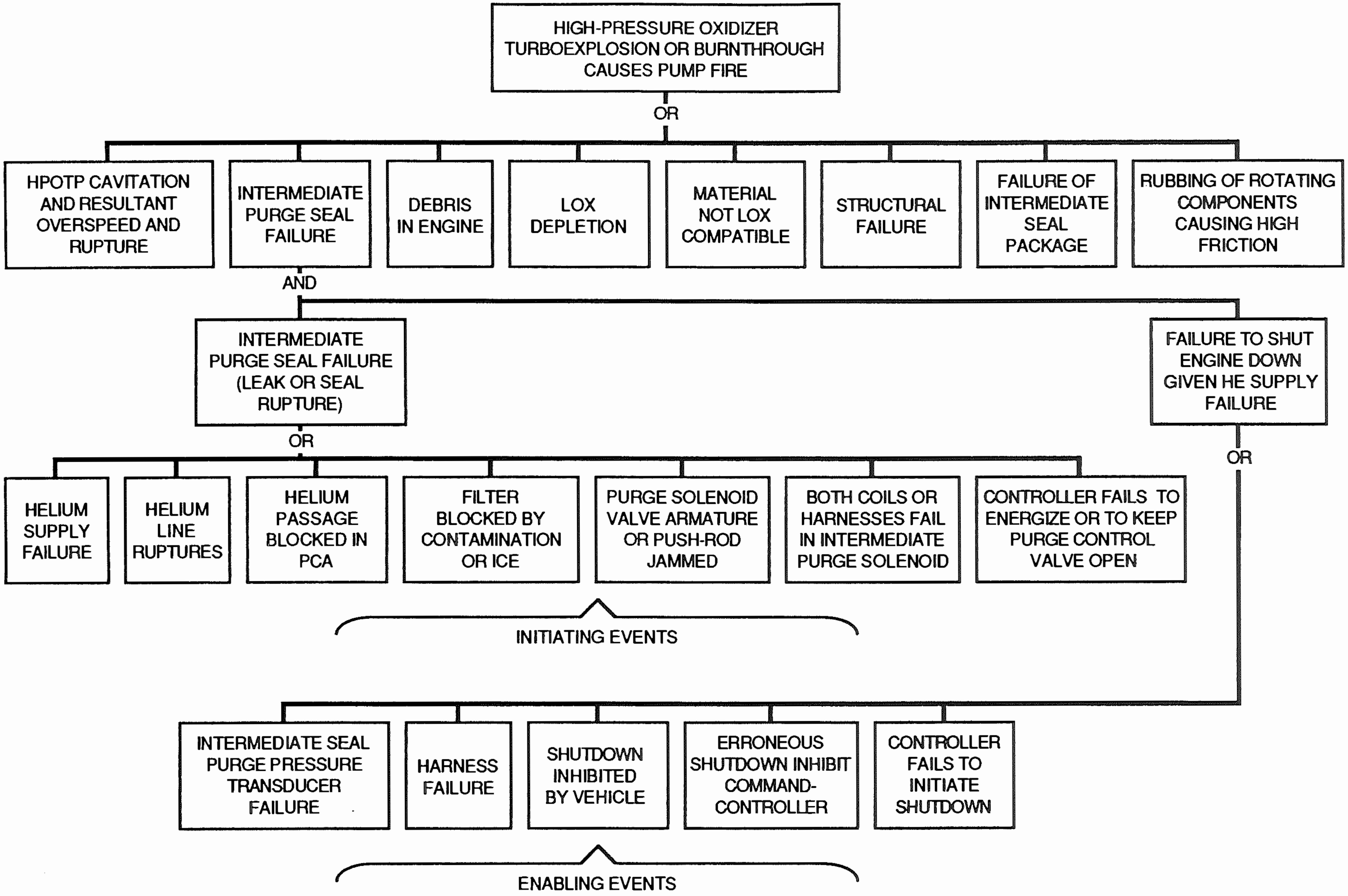

HIGH-PRESSURE OXIDIZER TURBOEXPLOSION OR BURNTHROUGH CAUSES PUMP FIRE
OR
HPOTP CAVITATION AND RESULTANT OVERSPEED AND RUPTURE
INTERMEDIATE PURGE SEAL FAILURE
DEBRIS IN ENGINE
LOX DEPLETION
MATERIAL NOT LOX COMPATIBLE
STRUCTURAL FAILURE
FAILURE OF INTERMEDIATE SEAL PACKAGE
RUBBING OF ROTATING COMPONENTS CAUSING HIGH FRICTION
AND
INTERMEDIATE PURGE SEAL FAILURE (LEAK OR SEAL RUPTURE)
FAILURE TO SHUT ENGINE DOWN GIVEN HE SUPPLY FAILURE
OR
OR
HELIUM SUPPLY FAILURE
HELIUM LINE RUPTURES
HELIUM PASSAGE BLOCKED IN PCA
FILTER BLOCKED BY CONTAMINATION OR ICE
PURGE SOLENOID VALVE ARMATURE OR PUSH-ROD JAMMED
BOTH COILS OR HARNESSES FAIL IN INTERMEDIATE PURGE SOLENOID
CONTROLLER FAILS TO ENERGIZE OR TO KEEP PURGE CONTROL VALVE OPEN
INITIATING EVENTS
INTERMEDIATE SEAL PURGE PRESSURE TRANSDUCER FAILURE
HARNESS FAILURE
SHUTDOWN INHIBITED BY VEHICLE
ERRONEOUS SHUTDOWN INHIBIT COMMAND-CONTROLLER
CONTROLLER FAILS TO INITIATE SHUTDOWN
ENABLING EVENTS

FIGURE 4-20. FAULT TREE FOR HPOTP EXPLOSION OR BURNTHROUGH

Failure to shut engine down given a helium supply failure (the right-hand input to the AND gate in Figure 4-20) is an enabling event. There are five enabling subevents:

- intermediate seal purge pressure transducer failure
- harness failure
- shutdown inhibited by vehicle
- erroneous shutdown inhibit command by controller
- controller failure to initiate shutdown.

An important insight is that the failure of the controller appears as both an initiating event (seventh event in the first list) and enabling event (fourth or fifth event in the second list).

Hence, the failure of the controller can be a common cause initiating event (also called a special initiator.) It can cause a need for system shutdown and simultaneously fail the system shutdown mitigative features. The controller is seen to be a critical item with respect to safety. However, the controller is a dual channel device with dual power supplies. There would have to be multiple failures within the controller to cause seal failure and engine shutdown failure.

Other examples of the importance of identifying initiating and enabling events are given by the fault trees for improper ullage (air space) pressure in the liquid oxygen ($L0_2$) tank located in the external tank of the Space Shuttle. Scenarios are considered below which result in either tank overpressurization or underpressurization. In either case, loss of mission and crew results. There are three parallel flow control values (for each engine) that regulate the flow of gaseous oxygen (GO_2) from the HPOTP to the $L0_2$ tank. Each control valve has a dedicated gauge type sensor. The gauge transducers have individual ambient ports that fail due to plugging from contamination or icing and could result in a transducer reading low. Failure of one sensor reading lower than actual tank pressure will open the flow control valve early. Tank pressure will remain within nominal limits with one sensor failing low. If two or three sensors read lower than actual pressure and the vent/relief valve fails closed, tank overpressurization will result. Two or more sensors reading higher than actual pressure will cause flow control valves to shut off too soon causing tank underpressurization, resulting in insufficient net positive suction pressure, subsequent pump cavitation, and explosion.

In either scenario, the failure of the first sensor is the enabling event, failure of the second sensor is the initiating event. Also, in the case of tank overpressurization, failure of the relief valve in the stuck closed position is an enabling event.

4.9 QUANTITATIVE RISK ASSESSMENT: DATA ANALYSIS

It is to be noted that expression 1 in Table 4-3 was in an exact form for the computation of accident frequency. In general, the occurrence of the Top Event (or more generally, the occurrence of an accident) is modeled as follows:

- initiating event occurs
- system is in a critical system state for the occurrence of the initiating event.

Since initiating events place demands on system mitigative features to respond, two quantities are of interest from a reliability viewpoint:

- initiating event occurrence frequency
- probability that system mitigative features fail to operate when the initiating event occurs (enabling event unavailability).

To compute these quantities, one must know the maintenance policies to which system components are subjected. For reliable systems, and assuming an exponential failure distribution, the component failure rate, λ, the conditional probability of failure per unit time, is an accurate approximation to the failure frequency, $w_f(t)$. λ dt is the probability of failure in [t,t+dt] given no failure in [0,t]; $w_f(t)dt$ is the probability of failure in [t,t+dt] given no failure at time t. It is to be kept in mind that an initiating event can occur more than once; e.g., a valve that cycles three times during a mission and can cause a catastrophic system failure if enabling conditions preexist when the valve cycles.

Enabling event unavailability, q, is a function of the following reliability parameters:

- the component mean-time-between failures, μ ($1/\lambda$, if λ is constant)
- inspection interval, θ
- mean repair or restoration time, τ

From an importance ranking viewpoint, it is to be noted that changes in these parameters can affect component unavailability and hence accident frequency. Figure 4-21 displays component unavailability and failure frequency for the following maintenance policies:

- no repair
- repair, announced failure
- repair, unannounced failure.

No repair means that when failure occurs, it is not detected for the remainder of the system life; e.g., a plug in a wet pipe sprinkler system. An announced failure, also called a revealed failure, is known at the time of the failure. An unannounced failure, also called a latent failure, is not detected until the end of some inspection interval. It is assumed in Figure 4-21 that λ is constant. For all three maintenance policies, λ is an accurate upper bound approximation to $w_f(t)$ for reliable systems. See Apostolakis (1974) and Lambert (1975) for a more detailed description of maintenance policies and their significance in risk analysis.

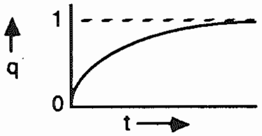
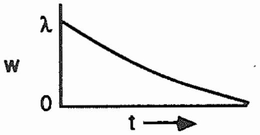
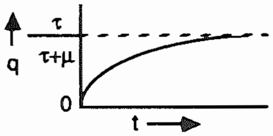
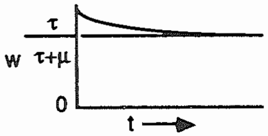
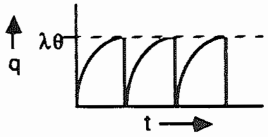
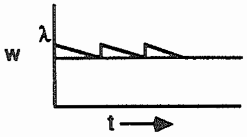
In modeling operator recovery, human errors and predictions of the probabilities of their occurrence are considered. Consult Swain et al. (1983) and Hannaman et al. (1984) for a discussion of human reliability assessment; see also Section 6.6, below.

Table 4-5 lists the values for the basic event parameters of concern in the analysis of the pressure tank system. Where there are little or no application-specific data available to support the evaluation of the basic event parameters, generic sources of data may be usable. Refer to the Air Force Rome Air Development Center (1975) for generic sources of data for the aerospace industry.

A comprehensive discussion of data development procedures for risk assessment is provided in Chapter 6, below.

4.10 ACCIDENT FREQUENCY EXPRESSION

In this section is discussed the modeling of the expected number of times per unit time that the system makes a transition from the unfailed to the failed state as defined by a Top Event in a fault tree.

MAINTENANCE POLICY	COMPONENT UNAVAILABILITY*	ASYMPTOTIC VALUE	COMPONENT UNAVAILABILITY VERSUS TIME	COMPONENT FAILURE FREQUENCY	ASYMPTOTIC VALUE	COMPONENT FAILURE FREQUENCY VERSUS TIME
1. No Repair	$1\text{-exp}(-\lambda t) \leq \lambda t$	1		$\lambda \exp(-\lambda t)$	0	
2. Repair Announced Failure	$\frac{\tau}{\mu+\tau}\left[1\text{-exp}\left[\left(-\frac{\mu+\tau}{\mu\tau}\right)t\right]\right.$	$\frac{\tau}{\mu+\tau} \leq \lambda t$		$\frac{1}{\mu+\tau}\left[1\text{-}\frac{\tau}{\mu}\exp\left(-\frac{\mu+\tau}{\mu\tau}t\right)\right]$	$\frac{1}{\mu+\tau} < \lambda$	
3. Repair Unannounced Failure	$1\text{-exp}[-\lambda(t-(n-1)\theta)]$ $(n-1)\,\theta \leq t \leq n\,\theta$ $n = 1,2,3,\ldots$	$\frac{\lambda\theta}{2} + \frac{\tau}{\tau+\theta}$ (Average Unavailability)		$\lambda \exp[-\lambda(t-(n-1)\theta)]$ $(n-1)\,\theta \leq t \leq n\,\theta$ $n = 1, 2, 3, \ldots$	$\lambda \exp\left[-\frac{\lambda\theta}{2}\right] < \lambda$	

* μ = Mean-Time-Between-Failures
τ = Mean-Time-To-Restore
θ = Scheduled Inspection Interval

FIGURE 4-21. COMPONENT UNAVAILABILITY AND FAILURE FREQUENCY - CONSTANT λ

Component Failure Mode	Basic Event Type	Failure Rate, λ, or Unavailability, q
Tank Rupture Under Normal Load, PT	Initiator	$\lambda_{PT} = 10^{-8}$ /cycle
Timer Contacts Fail to Open, T	Initiator	$\lambda_{T} = 10^{-4}$ /cycle
Voltage Surge, VS	Initiator	$\lambda_{VS} = 10^{-8}$ /cycle
Relief Valve Fails to Operate, R	Enabler	$\lambda_{R} = 3 \times 10^{-4}$ /hr $\theta_{R} = 1$ yr $q_{R} = 0.65$ (1) (2)
No or Slow Operator Response, O	Enabler	$q_{O} = 10^{-2}$ /demand
Manual Switch Fails to Open, S	Enabler	$q_{S} = 10^{-4}$ /demand
Pressure Gauge Stuck, G	Enabler	$q_{G} = 10^{-5}$ /hour

Note: (1) It is assumed that $\tau << \theta$.

(2) Expression $q = 1 - \{1 - \exp(-\lambda\theta)\}/\lambda\theta$ is used since $\lambda\theta$ is not small.

TABLE 4-5. DATA FOR PRESSURE TANK BASIC EVENTS

If it is assumed that the probability of occurrence of two initiating events in a differential time is zero (as is always the case for reliable systems), then the Top Event occurrence frequency, W(t), is the sum of the frequencies at which initiating events cause system failure; i.e.,

$$W(t) = \sum_{i=1}^{n} \Pr \left\{ \text{System is in a critical system state for the occurrence of initiating event i} \right\} * w_{f,i\,(t)}$$

$$= \sum_{i=1}^{n} \Pr \left\{ {}_{k}U_{(i)} E_{i,k} \right\} * w_{f,i\,(t)} \tag{4-2}$$

where

$\Pr$ = probability
$E_{i,k}$ = event that mincutset k containing initiating event i occurs (with event i set equal to true)
${}_{k}U_{(i)}$ = Boolean union of mincutsets k containing initiating event i, i = 1, ..., n
$w_{f,i\,(t)}$ = frequency of initiating failure event i

To evaluate the terms in parentheses in equation (4-2), assumptions must be made with regard to statistical dependency of component failures. One usually can make the following assumptions:

- System is reliable (i.e., the probability of the simultaneous occurrence of two or more mincutsets is small. This assumption is called the rare event approximation.)
- Basic events are statistically independent
- λ, the conditional probability of failure per unit time, is an accurate approximation for failure frequency.

The following notation is used:

i is an index for initiating events
j is an index for mincutsets, K
l is an index for enabling events
q denotes enabling event probability
ε means "belongs to"

Equation (4-2) then becomes

$$W(t) = \sum_{i=1}^{n} \{ \sum_{\substack{j \\ \text{such that} \\ i \varepsilon K_j}} \prod_{\substack{l \\ \text{such that} \\ l \varepsilon K_j \\ l \neq i}} q_l \} \lambda_i \qquad (4\text{-}3)$$

The term in brackets in equation (4-3) is a first order approximation for the critical state unavailability for initiating event i. The inner sum in equation (4-3) is the sum of all the mincutset frequencies containing initiating event i. Expanding equation (4-3) gives the sum of all mincutset frequencies and is a first order approximation. Equation (4-3) is generally sufficiently accurate for most risk calculations.

For the pressure tank system, equation (4-3) becomes (see again Table 4-5 for notation):

$$w(t) = \lambda_{PT} + q_R \lambda_{VS} + \{q_R\, q_G + q_R\, q_O + q_R\, q_S\}\, \lambda_T \qquad (4\text{-}4)$$

$$= \lambda_{PT} + q_R \lambda_{VS} + q_R\, \{q_G + q_O + q_S\}\, \lambda_T \qquad (4\text{-}5)$$

$$= \lambda_{PT} + q_R \lambda_{VS} + q_R\, q_G\, \lambda_T + q_R\, q_O\, \lambda_T + q_R\, q_S\, \lambda_T \qquad (4\text{-}6)$$

The expansion of equation (4-4) to the sum of the mincutset frequencies is thus given by equation (4-6).

If it is assumed that there is on the average one operating cycle per hour and if the basic event data in Table 4-5 are used, equation (4-4) gives

$$\begin{aligned} W(t) &= 1.0X10^{-8} + 0.65X10^{-8} + 0.65\ (1.0X10^{-5} + 1.0X10^{-2} + 1.0X10^{-4})1.0X10^{-4}/hr \\ &= 6.7 \times 10^{-7}/hr \text{ (or per cycle)} \\ &= 5.9 \times 10^{-3}/yr \qquad (4\text{-}7) \end{aligned}$$

The mean time to the occurrence of the Top Event is the reciprocal of W(t), i.e., 170 years.

Thus far, only basic failure rates and repair rates that are constant in time have been considered. In some situations, failure rates may exhibit a burn-in and/or a wear-out characteristic. In this case, a multi-parameter probability distribution for time to failure can be used, such as a gamma or Weibull distribution. As discussed in Lambert (1975), it is a straightforward procedure to include such distributions in reliability calculations. In fault trees, the Top Event occurrence frequency is then also not constant in time and must be integrated over time to obtain the expected number of occurrences of the Top Event per unit time.

4.11 IMPORTANCE EXPRESSIONS

The development of importance expressions for components or systems is carried out in three basic steps:

- formation of a new Top Event that is the Boolean union of the mincutsets containing either the initiating or enabling event
- use of equation (4-2) to compute the frequency of occurrence of the new Top Event (in initiating event importance expressions,only one event can function as the initiating event for the new Top Event)
- division of the results in the second step by the accident frequency.

Stated mathematically, the importance expression for basic events weighted according to accident frequency is

$$I_{AF} = \frac{\text{Frequency of the Boolean union of mincutsets containing the event of interest}}{\text{Top Event frequency, } W(t)} \qquad \textbf{(4-8)}$$

The above importance expression is simply the fractional contribution of the mincutsets containing either the initiating or the enabling event to the total accident frequency. If the first order approximation for the probability of a union is used, the numerator in expression (4-8) for initiating events becomes the sum of all the terms in the Top Event occurrence frequency expression, which contain the initiating failure frequency, λ ; for enabling events, the numerator is the sum of all terms containing the enabling event probability, q.

Table 4-6 lists the importance expressions and values for the basic events and system failure modes for the pressure tank system. The weighting is according to Top Event occurrence frequency, W(t). In this case, W(t) is constant. In many cases in risk assessment, W(t) is constant or can be satisfactorily represented by a constant. It is generally assumed that first order approximations are valid; i.e., that use of equation 4-3 results in an accurate calculation. Examining the expressions in Table 4-6, one sees that for initiating events, the numerator is a linear function of the failure frequency; for enabling events, the numerator is a linear function of the enabling event unavailability. Conceptually, enabling event importance is a contributory measure of importance since enabling events do not cause an accident to occur.

In Table 4-6 it is seen finally that the following events have the highest importance values:

- timer contacts fail to open
- relief valve fails to operate
- no or slow operator response.

The rupture of the tank under load (a single event mincutset) has probabilistic importance of order 10^{-2} and, despite its greater structural importance, its low probability of occurrence results in its quantitative importance being less significant than the events listed above.

4.12 RISK MITIGATIONS EVALUATION

To mitigate the effects of the failures described in the previous section, one can, for instance, incorporate changes in reliability parameters of the involved components, e.g., one can incorporate the following improvements or mitigations (new values of reliability parameters are indicated in parentheses):

- Install a timer that fails less frequently (a failure rate of 1.0 x 10^{-5} per cycle)
- Employ special operator procedures (operator failure probability of 1.0 x 10^{-3})
- Inspect the relief valve and manual switch before each operating cycle
- Install an identical timer in series with the first one (inspect each timer once a month)

Component Failure Mode or System Failure Mode	Mathematical Expression (1)	Value
Pressure Tank Rupture Under Load, PT	$\lambda_{PT}/W(t)$	1.5×10^{-2}
Timer Contacts Fail to Open, T	$q_R\left\{q_G + q_O + q_S\right\}\lambda_T/W(t)$	0.97
Voltage Surge, VS	$q_R\lambda_{VS}/W(t)$	9.7×10^{-3}
Relief Valve Fails to Operate, R, or Pressure Protection Fails, PP	$(q_R\lambda_{VS} + q_R\left\{q_G + q_O + q_S\right\}\lambda_T)/W(t)$ $= \left\{q_{PP}\lambda_{VS} + q_{PP}q_{OS}\lambda_T\right\}/W(t)$	0.98
No or Slow Operator Response, O	$q_Rq_O\lambda_T/W(t)$	0.97
Manual Switch Fails to Open, S	$q_Rq_S\lambda_T/W(t)$	9.7×10^{-3}
Pressure Gauge Stuck, G	$q_Rq_G\lambda_T/W(t)$	9.7×10^{-4}
Operator Shutdown System Fails, OS	$q_R\left\{q_G + q_O + q_S\right\}\lambda_T/W(t)$ $= q_{PP}q_{OS}\lambda_T/W(t)$	0.98

Note : (1) $W(t) = \lambda_{PT} + q_R\lambda_{VS} + (q_R\left\{q_G + q_O + q_S\right\})\lambda_T$

TABLE 4-6. IMPORTANCE RANKINGS FOR PRESSURE TANK SYSTEM

Table 4-7 displays the reduction in risk when the above mitigations are incorporated. The third improvement provides the greatest risk reduction (98%). For the first three improvements listed above, only a component's reliability parameter changes. For the third improvement, it is assumed that the operability of each component prior to the operating cycle is checked, so that there are no pre-existing failures at the start of the operating cycle.

Potential Improvement (Change in reliability parameter)	W(new) [per hr]	Risk Reduction Ratio, W(new)/W(old) (1)
Install a more reliable timer (failure rate decreases by a factor of 10)	8.7×10^{-8}	.13
Special operator procedures (failure probability decreases by a factor of 10)	8.7×10^{-8}	.13
Inspect relief valve and manual switch before each cycle	1.0×10^{-8}	.02
Incorporate a redundant timer in series with the first (inspection interval of one month)	6.7×10^{-8}	.10
Note: (1) W(old) = 6.7×10^{-7}		

Table 4-7. Effects of Potential Improvements in the Pressure Tank System

The fault tree logic changes when the fourth improvement is introduced and so a new initiating event fault tree must be generated. In this case, there is functional redundancy in preventing the occurrence of the initiating event. Mincutsets 3, 4, and 5 in Table 4-8 can have two basic events that can act as an initiating event. In other words, there are two time sequences implied by one mincutset. As was described in Section 4.8.1, mincutsets defining conditions for fire and explosion have this property.

Min Cut Set	Description
1	• Tank Ruptures Under Normal Load (i)
2	• Voltage Surge (i) • Relief Valve Fails to Operate (e)
3	• Timer Contacts 1 Fail to Open (i) • Timer Contacts 2 Fail to Open (i) • Relief Valve Fails to Operate (e) • Pressure Gauge Stuck (e)
4	• Timer Contacts 1 Fail to Open (i) • Timer Contacts 2 Fail to Open (i) • Relief Valve Fails to Operate (e) • No or Slow Operator Response (e)
5	• Timer Contacts 1 Fail to Open (i) • Timer Contacts 2 Fail to Open (i) • Relief Valve Fails to Operate (e • Manual Switch Fails to Open (e)

Notes: (i) Denotes an initiating event
(e) Denotes an enabling event

Table 4-8. Listing of Mincutsets for Pressure Tank System with Two Timers

One can still compute a new frequency for the random failure of both timers, timer 1 and timer 2, as

$$q_1\lambda_2 + q_2\lambda_1 = 2q_T\lambda_T \qquad (4\text{-}9)$$

where

$$\begin{aligned} q_T &= \lambda_T \, \theta_T \times 12 \\ &= 1 \times 10^{-4} \times 720/2 \\ &= 3.6 \times 10^{-2} \end{aligned}$$

Hence, when a second timer is placed in series, the effective failure frequency is reduced by a factor of $2q_T$.

When considering the random failure of both timers, it is again to be recalled that the timer that fails first does not cause pump overrun; it is the second timer failure that causes pump overrun. Expression (4-8) represents two possible sequences of events. In the first term, timer 1 fails first (it is the enabling event) and timer 2 fails second (it is the initiating event). A reverse ordering of events applies in the second term.

For the Space Shuttle, there is redundancy in the number of main engines. A non-catastrophic shutdown of one of the three engines will not result (directly) in loss of vehicle or crew. Shutdown of two engines can result in a loss of vehicle and crew. In this case, there are six possible time sequences of pairs of engine failures to consider; i.e., (1, 2), (2, 1), (1, 3), (3, 1), (2, 3) and (3, 2), where 1 denotes the failure of engine 1, etc.

4.13 ASSUMPTIONS AND SENSITIVITY ANALYSIS

In risk assessment, it is common to assess the effects of various assumptions made in the analysis, particularly when issues involving human action and recovery are addressed. In addition, one makes assumptions about pre-existing conditions prior to system operation.

For example, in Section 3.3 the assumption was made that the pressure tank starts each cycle unpressurized; i.e., it was assumed that the operator opens the discharge valve after each cycle. However, the pressure tank can start the cycle pressurized if the operator fails to open the discharge valve from the previous cycle. In this case, if he erroneously resets the timer and then fails to shut the system down, overpressure sufficient to rupture the tank will occur if the relief valve fails to operate.

It is seen that this sequence is dominated by a chain of human error events. One can assess the impact of this sequence by constructing a new fault tree or event tree and then calculating a new accident frequency (consult Swain and Guttmann, 1983, and Hannaman et al., 1984).

As another example of sensitivity to assumptions, consider the detection system that detects hardware failures within a system. The detection system generally consists of software, hardware, and human elements.

A common assumption made in risk assessment is that perfect inspection occurs. More realistically, imperfect inspection can affect the fault duration of events and cause an increase in the accident frequency. For example, if the operator fails to observe nitrogen purge pressure and flammable gas concentration, then the fault duration times for conditions that can permit fire or explosion increase, resulting in an increase in the probability of fire or explosion.

Imperfect inspection can also result from hardware failures (Hasegawa et al., 1979). Studies conducted by Lawrence Livermore National Laboratory (LLNL) and British Gas Corporation (Morgan and Andrews, 1984), indicate that the reliability of the inspection system significantly affects the availability of fire protection systems. In the LLNL study, a zone indicating unit (ZIU) has a supervisory circuit that detects failure of electronic components within the ZIU. If the supervisory circuit fails, then these failures are not detected until the end of the inspection interval, which is three months. As shown in the fault tree in Figure 4-22, a component failure may either be announced (i.e., detected and identified) or unannounced depending upon the operability of the detection circuit. The dominant mincutsets in the LLNL study were of the following description: trouble light fails off, and an electrical component within the ZIU fails open-circuit.

4.14 RISK ANALYSIS COMPUTER CODES

Available computer codes for risk analysis are of three types:

- qualitative analysis (Table 4-9)
- quantitative analysis (Table 4-10)
- uncertainty analysis (Table 4-11)

Tables 4-9, 4-10, and 4-11 are taken from the PRA Procedures Guide (Nuclear Regulatory Commission, 1983). In addition, there are several other commercially available PRA packages that are described below.

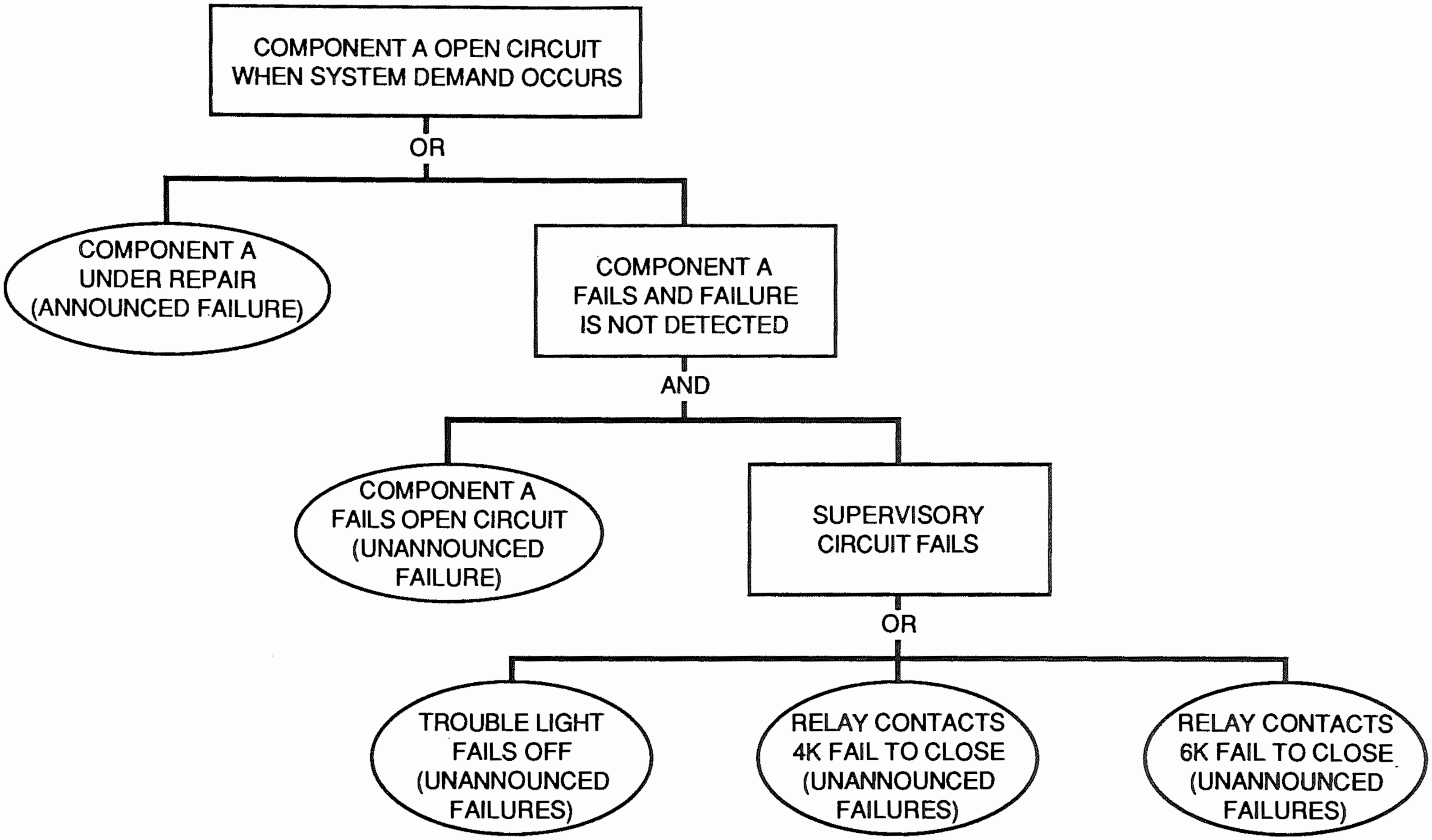

COMPONENT A OPEN CIRCUIT WHEN SYSTEM DEMAND OCCURS
OR
COMPONENT A UNDER REPAIR (ANNOUNCED FAILURE)
COMPONENT A FAILS AND FAILURE IS NOT DETECTED
AND
COMPONENT A FAILS OPEN CIRCUIT (UNANNOUNCED FAILURE)
SUPERVISORY CIRCUIT FAILS
OR
TROUBLE LIGHT FAILS OFF (UNANNOUNCED FAILURES)
RELAY CONTACTS 4K FAIL TO CLOSE (UNANNOUNCED FAILURES)
RELAY CONTACTS 6K FAIL TO CLOSE (UNANNOUNCED FAILURES)

FIGURE 4-22. GENERIC FAULT TREE

Table 4-9. Computer Codes for Qualitative Analysis[a] (Page 1 of 2)

Code	Input	Limit on number of gates or events	Types of gates	Limit on number or size of cut sets[b]	Method of generating cut sets[a]	Other outputs	Fault-tree truncation	Other features	Type of computer, language, and availability
ALLCUTS	8-character alphanumeric names, control information, primary-event probability, fault-tree description	175 primary events and 425 gates	AND OR	Up to 1000 cut sets can be generated	Top-down successive Boolean substitution	Cut sets in specified probability range, cut set and top-event probability	Minimal cut sets, probability	Fault-tree plotting option	IBM 360/370 CDC 7600 Fortran IV
FATRAN	8-character alphanumeric names, control information, fault-tree description	None	AND OR	None	Top-down successive substitution with gate-coalescing option	Minimal cut sets up to specified order	Minimal cut sets	—	CDC Cyber 76 Fortran IV Available from EG&G Idaho, Inc.
FTAP	8-character alphanumeric names, control information, fault-tree description	None; computer memory is limiting factor	AND OR K-of-M NOT	Minimal cut sets of up to order 10 can be generated	Top-down, bottom-up, and Nelson method (prime implicants)	Minimal cut sets and prime implicants	Minimal cut sets	Independent subtrees automatically found and replaced by module	IBM 360/370 CDC 6600-7600 Fortran IV Available from Operations Research Center, University of California, Berkeley
MOCUS	8-character alphanumeric names, control information, fault-tree description	None	AND OR INHIBIT	Minimal cut sets of up to order 20 can be generated	Top-down successive Boolean substitution	Path sets	Minimal cut sets	Cut sets can be automatically punched in cards or on-line data sets for use by KITT or SUPERPOCUS	IBM 360/370 CDC 7600 Fortran IV Available from Argonne Software Center
PL-NOO	79-character alphanumeric names, control information, fault-tree description failure data	None; computer memory is limiting factor	AND OR NOT K-of-N	None	Bottom-up modularization and decomposition of fault tree into best modular representation	Probability of top event, time-dependent characteristics of top event, minimal cut sets, uncertainty for top event	Minimal cut sets	Option of not generating minimal cut sets for quantifying fault trees	IBM 360/370 PL/I Available from Argonne Software Center
PREP	8-character alphanumeric names, control information, fault-tree description	2000 primary events and 2000 gates	AND OR INHIBIT	Minimal cut sets of up to order 10 can be generated	Combinatorial testing	No	Minimal cut	IBM 360/370 sets can be automatically punched on cards or on-line data sets for use in KITT or SUPERPOCUS	CDC 7600 Fortran IV Available from Argonne Software Center

Table 4-9. Computer Codes for Qualitative Analysis[a] (Page 2 of 2)

Code	Input	Limit on number of gates or events	Types of gates	Limit on number or size of cut sets[b]	Method of generating cut sets[a]	Other outputs	Fault-tree truncation	Other features	Type of computer, language, and availability
SETS	16-character alphanumeric names, user's program, failure data, fault-tree description	8000 events (gates and primary events together)	AND OR INHIBIT PRIORITY Exclusive or special	None	Top-down Boolean substitution, but user's program can be designed for any other method	Probability of minimal cut sets, prime implicants	Yes, based on both cutset order and probability	Automatic fault-tree merging and plotting; on-line data sets can be stored on tapes for use in other runs; independent subtrees can be obtained to simplify cut-set generation	CDC 7600 Fortran IV Available from Argonne Software Center
SIFTA	10-character alphanumeric names, control information, failure data, fault-tree description	None; computer memory is limiting factor	AND OR K-of-N	No cut sets generated	Pattern-recognition technique to reduce structure of tree; numerical simulation to calculate probabilities	New structure of tree after reduction; probability of top event	Independent branches of tree with small probability	Handles trees with multiple top events; merging of fault trees possible; fault trees can be plotted	HP-1000 Available from Atomic Energy Control Board, Ottawa, Canada
TREEL and MICSUP	8-character alphanumeric names, control information, fault-tree description	None; computer memory is limiting factor	AND OR INHIBIT	Minimal cut sets of up to order 10 are generated	Top-down successive Boolean substitution	Path sets	Minimal cut sets	Can determine minimal sets of intermediate gates	CDC 6400 Fortran IV Available from Operations Research Center, University of California, Berkeley
WAMCUT WAMCUT II	10-character alphanumeric names, control information, failure data, fault-tree description Center	1500 primary events and 1500 gates	AND OR NOT NOR NAND ANOT ONOT K-of-N	Up to 2000 minimal cut sets of any order can be generated	Bottom-up Boolean substitution; WAMCUT-II finds independent subtrees, replaces them by pseudocomponents, then uses top-down Boolean substitution	Probabilities of minimal cut sets and top event, first and second moments of minimal cut sets and top event	Yes, based on both cut-set order and probability	Plot option, can generate minimal cut sets of intermediate gates	CDC 7600 IBM 370 Extended Fortran IV Available from EPRI Code

[a] All the codes listed here have routines for checking input errors. These routines are very extensive in the codes FTAP, MOCUS, PREP, SETS, SIFTA, TPEEL-MISCUP, and WARCUT. ALLCUTS uses the auxiliary code BRANCH for checking input errors.

[b] Or prime implicants.

Table 4-10. Computer Codes for Quantitative Analysis

Code	Input	Quantitative calculations	Importance calculation	Other features	Type of computer and availability[a]
FRANTIC, FRANTIC II	Reduced system equation or minimal cut sets, primary-event failure data	Time-dependent calculation; nonrepairable, monitored, and periodically tested primary events are handled; uncertainty analysis for failure rates in conjunction with time-dependent calculation	No	Can model human-error and dependent-failure contributions; FRANTIC II can handle time-dependent failure rates and incorporates effect of renewal on aging	IBM 360/370 Available from Argonne Software Center
GO	GO chart[b] and fault-tree failure data	Only time-independent calculations for gates and top event; nonrepairable or periodically tested primary events are handled	No	Cut sets for selected gates and probability truncation of cut sets up to order 4	CDC 7600 Available from EPRI Code Center
ICARUS	Reduced system equation, choice of testing scheme, failure data	Average unavailability, optimal test interval, relative contributions of testing, repair, and random failures	No	Three testing schemes available: random testing, uniformly staggered testing, and nearly simultaneous testing	IBM 360/370 Available from Argonne Software Center

Table 4-10. Computer Codes for Quantitative Analysis (Continued)

Code	Input	Quantitative calculations	Importance calculation	Other features	Type of computer and availability[a]
IMPORTANCE	Minimal cut sets, primary-event failure data	Top-event point-estimate probability or unavailability	Can calculate the following: Birnbaum, criticality, upgrading function, Fussell-Vesely, Barlow-Proschan, steady-state Barlow-Proschan, sequential contributory	Can rank cut sets and primary events on basis of each importance measure	CDC 7600 Available from Argonne Software Center
KITT-1, KITT-2	Minimal cut sets supplied directly or by MOCUS or PREP; primary-event failure data	Time-dependent un-availability for primary events, minimal cut sets, and top event; failure rate, expected number of failures, and unreliability for top event and minimal cut sets	Fussell-Vesely importance calculations for primary events and minimal cut sets	KITT-2 allows each component to have unique time phases and thus failure and repair to vary from phase to phase	IBM 360/370 CDC 7600 Available from Argonne Software Center

Table 4-10. Computer Codes for Quantitative Analysis (Continued)

Code	Input	Quantitative calculations	Importance calculation	Other features	Type of computer and availability[a]
RALLY	Fault-tree description, control information, failure data	Average unavailabilities and failure frequencies for top event; time-dependent calculation possible through use of minimal cut sets; uncertainty analysis possible by using minimal cut sets; normal, lognormal, Johnson, extreme value-1, Weibull, gamma, and exponential distributions are handled	Code CRESSEX in RALLY can perform importance calculations	Can handle up to 1500 components and 2000 gates; can determine minimal cut sets using either a simulative or analytical way	IBM 360/370
RAS	Fault-tree description or minimal cut sets; failure and repair rates	Time-dependent unavailability, expected number of failures, and frequency of top event	No	Phased-mission analysis possible, if fault tree is input, minimal cut sets will be calculated	CDC 7600 Available from Argonne Software Center
SUPERPOCUS	Minimal cut sets, component failure data, time at which calculations are performed	Time-dependent unavailability, reliability, and expected number of failures for minimal cut sets and top event	Yes	Ranks minimal cut sets on basis of importance; can read cut sets directly from MOCUS or PREP	IBM 360/370 CDC 7600 Available from Dept. of Nuclear Engineering, University of Tennessee
WAM-BAM	Fault-tree description, primary-event failure data	Point unavailability for top event and intermediate gates; no time-dependent analysis possible	No	Extensive error checking possible through WAM; probability truncation of fault tree; sensitivity analysis possible by WAM-TAP preprocessor instead of WAM	CDC 7600 Available from EPRI Code Center

[a] All the codes listed here are written in Fortran IV.

[b] A GO chart (see Section 4.6.3) is a chart that resembles a schematic of system primary events and their relations via a set of 16 Boolean operators.

Table 4-11. Computer Codes for Uncertainty Analysis

Code	Input	Method of uncertainty analysis	Type of statistical distribution	Other features	Type of computer and availability[a]
BOUNDS	Reduced system equation or minimal cut sets, primary-event failure data	Mathematical combination of uncertainties; output includes two moments of minimal cut sets and the top event	Johnson, empirical	Can handle multiple system functions with multiple data input descriptions; can fit Johnson-type distribution to the top event	IBM 360/370 Available from University of California at Los Angeles
MOCARS	Minimal cut sets or reduced system equation, primary-event failure data	Monte Carlo simulation	Exponential, Cauchy, Weibull, empirical, normal, lognormal, uniform	Microfilm plotting of output distribution; Kolmogorov-Smirnov goodness-of-fit test on output distribution is possible	CDC Cyber 76 Available from Argonne Software Center
PROSA-2	Reduced algebraic function for system representation, failure data	Monte Carlo simulation	Normal, lognormal, uniform, any distribution in the form of a histogram, truncated normal beta	Can correlate input parameters; no sorting necessary to obtain the top-event histogram	IBM 370 Available from Argonne Software Center
SAMPLE	Minimal cut sets or reduced system equation, primary-event failure data	Monte Carlo simulation	Uniform, normal, lognormal	Used in the Reactor Safety Study; output is a probability distribution for the top event	IBM 360/370 Available from Argonne Center
SPASM	Fault tree or reduced system equation, component-failure data	Mathematical combination (similar to BOUNDS)	Lognormal	Works in conjunction with WAMCUT	CDC 7600 Available from EPRI Code Center
STADIC-II	Reduced system equation, primary-event failure data	Monte Carlo simulation (similar to SAMPLE)	Normal, lognormal, log-uniform, tabular input distribution	Has a better and efficient method of sorting the probabilities obtained in each trial	PRIME, UNIVAC- 1180, Available from General Atomic Company

[a] All the codes listed here are written in Fortran IV.

Codes that perform uncertainty analysis calculate confidence intervals for the Top Event and accident frequency. There is statistical uncertainty in computation of accident frequency because there is statistical uncertainty in the basic event parameters such as:

- component failure rates
- component repair times
- maintenance frequency
- human error probability.

Uncertainty analysis can be conducted (1) by calculating the distribution of the Top Event analytically by the method of moments or (2) by Monte Carlo simulation. In some cases, a direct derivation is possible, at least approximately, of a Top Event's uncertainty distribution from the basic event uncertainty distributions (see Section 6.3).

4.14.1 NUSSAR-II, Safety and Reliability Analysis Software

NUSSAR-II is a safety and reliability software system with applications that include:

- fault tree construction and documentation
- system analyses
- human reliability analyses
- accident sequence analyses (through the use of mainframe fault tree computer codes)
- importance analyses
- Monte Carlo uncertainty analyses
- time-dependent reliability analyses

NUSSAR-II is completely menu-driven and supports the following main functions:

- The fault tree is displayed on the screen in full-color graphics. Fault tree models are constructed and edited using simple commands.

- Fault trees may be documented either as printouts or report-quality drawings. This requires a pen plotter or laser printer with Hewlett Packard Graphic Language (HPGL) capability.

- The user can create databases containing reliability data and descriptions for components/failure modes. As the user constructs the fault trees, the database automatically feeds the required data and descriptions to the basic events. NUSSAR-II instantaneously displays information on-screen.

- Fault trees are organized into frontline and support system models. NUSSAR-II permits the user to connect the logic between many fault trees so that the linked, interconnected fault tree may be analyzed.

- The user supplies the cutoff probability for the cut set calculation (below which cut sets are deemed insignificant), and NUSSAR-II evaluates the user's fault tree(s), providing the user with a list of minimal cut sets ranked by cut set probability. The calculated Top Event unavailability is exact - not based on the rare event approximation.

- NUSSAR-II features an on-screen cut set analyzer and editor to include common mode failures and recovery actions in the cut set lists. An additional analysis feature of this module includes the capability to change basic event data within the cut set list while observing the immediate impact on the Top Event unavailability.

- Cut set lists are processed through the importance analysis module to calculate various basic event importance measures, such as Fussell-Vesely, risk reduction worth, risk achievement worth, and criticality.

- To complete the analysis, the user can carry out a Monte Carlo simulation to derive the probability (i.e., uncertainty) distribution of the Top Event frequency or unavailability. NUSSAR-II analysis provides the user with the mean, variance, and different confidence levels of the Top Event probability distribution.

- Time dependent reliability calculations enable the user to assess the effects of component failure rates, repair times, and test intervals on system availability. The program permits the user to calculate the time-dependent evolution of system unavailability, conditional and unconditional failure intensities, expected number of failures, and the probability of at least one failure.

For more information, or to order a NUSSAR-II demonstration, call or write:

NUS Corporation
16835 West Bernardo Drive
Suite 202
San Diego, CA 92127
(616) 451-2131

4.14.2 SETS and Related Programs

The Set Equation Transformation System (SETS) is a computer program for the symbolic manipulation of Boolean equations. Symbolic manipulation changes equations from their original forms into more useful or desirable forms, particularly by the application of Boolean identities. The SETS program is an interpreter that reads, interprets, and executes SETS user programs. The user writes a SETS user program specifying the processing to be achieved and submits it, along with the required data, for execution by SETS.

SETS has been used for fault tree analysis, event tree analysis, vital area analysis, common cause analysis, and probabilistic risk assessment of nuclear reactor power plants. It has been used to verify circuit design implementation, determine minimum cost fire protection requirements for reactor plants, obtain solutions to combinatorial optimization problems with Boolean constraints, and determine the susceptibility of a facility to unauthorized access through nullification of sensors in its protection system.

Two auxiliary programs, the Set Evaluation Program (SEP) and the Fault Tree Drawing (FTD) Program, are used in conjunction with SETS. SEP is used to quantify fault tree minimal cutset equations produced by SETS. Equation probabilities can be computed by the rare-event approximation using point value component failure probabilities or by a sampling procedure using log-normal-distributed component failure probabilities. SEP also computes the Birnbaum and Fussell-Vesley importance measures. FTD is used to plot fault trees processed by SETS. Fault trees can be plotted either with or without event descriptions and the plot can be annotated with references and notes.

The SETS, SEP, and FTD programs are available on CDC and CRAY systems. An enhanced version of SETS and a Fault Tree Graphics (FTG) program to replace FTD will be available in early 1988 for CDC, CRAY, VAX, and IBM systems. FTG provides a graphics terminal capability for constructing and modifying fault trees processed by SETS and a hard-copy capability for producing annotated fault tree plots.

These computer programs and support and maintenance for them are available from:

R. B. Worrell
Logic Analysts
1717 Louisiana NE
Suite 202
Albuquerque, NM 87110

4.14.3 CAFTA+

CAFTA+ is a management tool to build, modify, and evaluate fault tree models. It includes a fault tree editor, a basic event data base, formatting routines, fault tree analysis routines, a cut set editor and fault tree plotting. The Fault Tree Editor is designed for both text editing for entering models, and logical editing for modifying trees. Syntax and logic checks are performed as models are built and modified. The Fault Tree Editor also links basic event reliability data from all trees into one base file. This Basic Event Data Base includes the basic event name, trees in which it is used, and information on how to calculate probability. The data base maintains a second data file that contains basic event types and failure modes.

Besides building fault tree models, CAFTA+ includes a fault tree and cut set processor, SAICUT. The fault tree processor can generate and quantify cut sets from fault tree models built by CAFTA+. Also, importance calculations can be performed on the cut sets stored and edited by the Cutset Editor. Since CAFTA+ may not be able to handle the analysis of large fault trees in a reasonable time, trees can be formatted and linked with reliability data for input into many analysis programs (e.g., WAM Series, SETS, FTAP). CAFTA+ can also download cut sets generated by the analysis programs for use in the Cutset Editor. Also, cut sets from independent subtrees can be used to replace subtrees in larger trees with single basic events, thereby reducing analysis time needed for larger trees.

Completed fault trees can be plotted on numerous devices including HP plotters and Apple Laser printers using SAIPLOT. CAFTA+ requires an IBM PC (XT/AT), or a compatible capability, with 640K and DOS 2.0 or later.

CAFTA is a software product of the Electric Power Research Institute (EPRI) developed under contract by Science Applications International Corporation (SAIC). CAFTA+ is a version of CAFTA enhanced by the fault tree and cut set processor SAICUT and the fault tree plotting program SAIPLOT.

For information contact:

Blake Putney or Jim Koren
SAIC
5150 El Camino Real
Suite C-31
Los Altos, CA 94022
(415) 960-3322

4.14.4 AutoCAD/TREEGEN Software Package

Erin Engineering has available a PC-based computer code package capable of conducting probabilistic risk assessment and fault tree analysis. The AutoCAD/TREEGEN software portion of the package allows the user to quickly prepare drafting-quality fault trees and event trees for use in PRA applications, as well as simplified Piping and Instrumentation (P&ID's) and electrical one-line drawings. In addition to providing the user with the drawing capabilities of the AutoCAD system (a commercial software package produced by Autodesk, Inc.), an additional program module allows the user to automatically generate input decks (for use with FTAP, POSTPROCESSOR, and IMPORTANCE codes, described below) directly from the information contained in the AutoCAD fault tree drawing files.

Figure 4-23 presents a flow chart of the overall execution sequence of the AutoCAD/ TREEGEN software package. As mentioned above, AutoCAD is used to generate the desired fault tree, event tree or PI&D drawings. ERIN Engineering and Research has created a library of custom-defined symbols and two customized on-screen menu formats to assist the user in creating these drawings.

If a fault tree drawing is being prepared, AutoCAD will create a fault tree data file in addition to the usual drawing file. This file can be used by TREEGEN and its associated programs to generate input decks for use by PRA analysis codes.

The code package includes:

- FTAP
- postprocessor to FTAP (optional use)
- IMPORTANCE
- MONTE

The inputs and outputs to each computer code are shown in Figure 4-24. FTAP performs the qualitative evaluation; IMPORTANCE, the probabilistic evaluation; and MONTE, the uncertainty analysis.

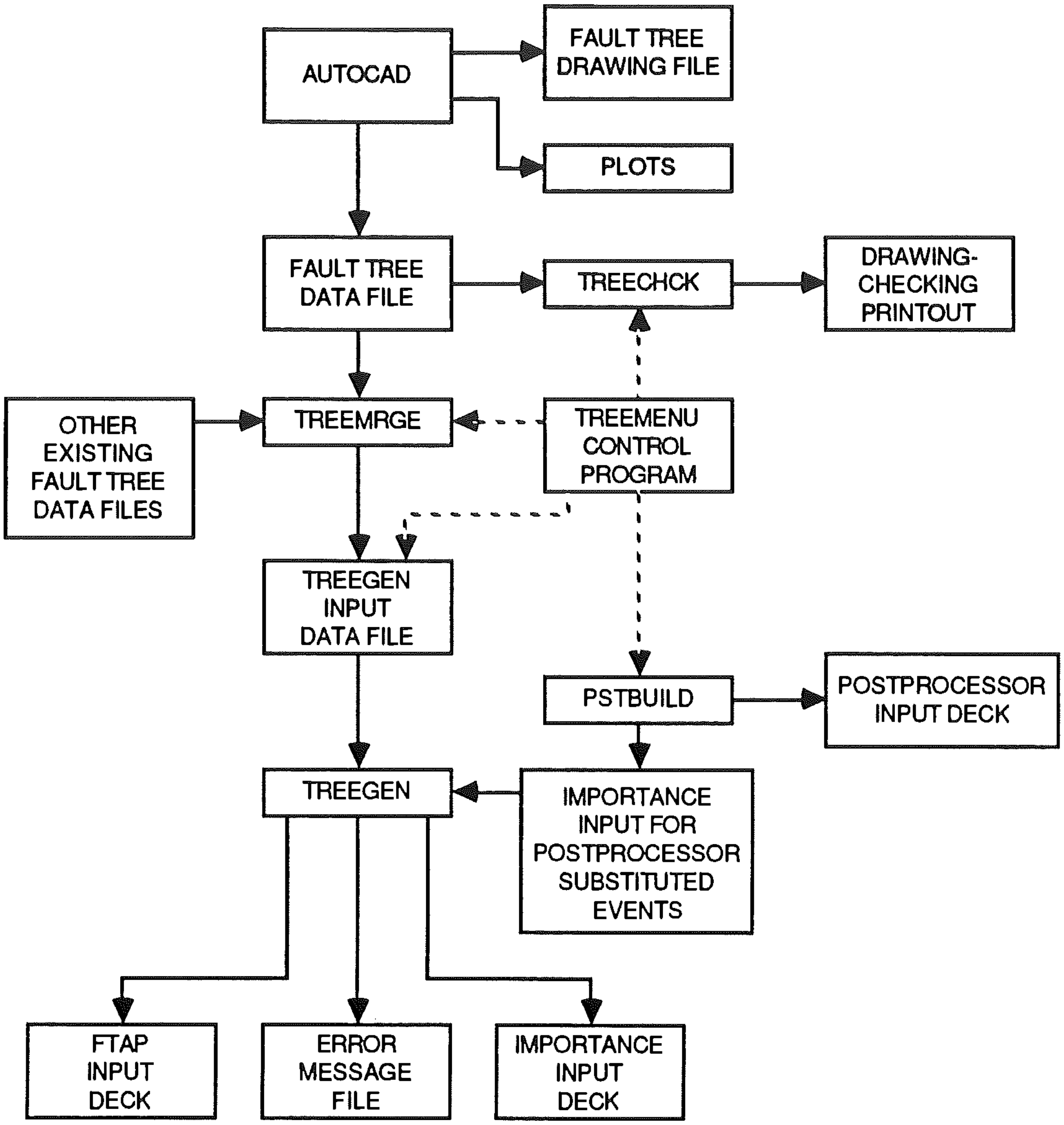


FIGURE 4-23. AUTOCAD/TREEGEN EXECUTION SEQUENCE

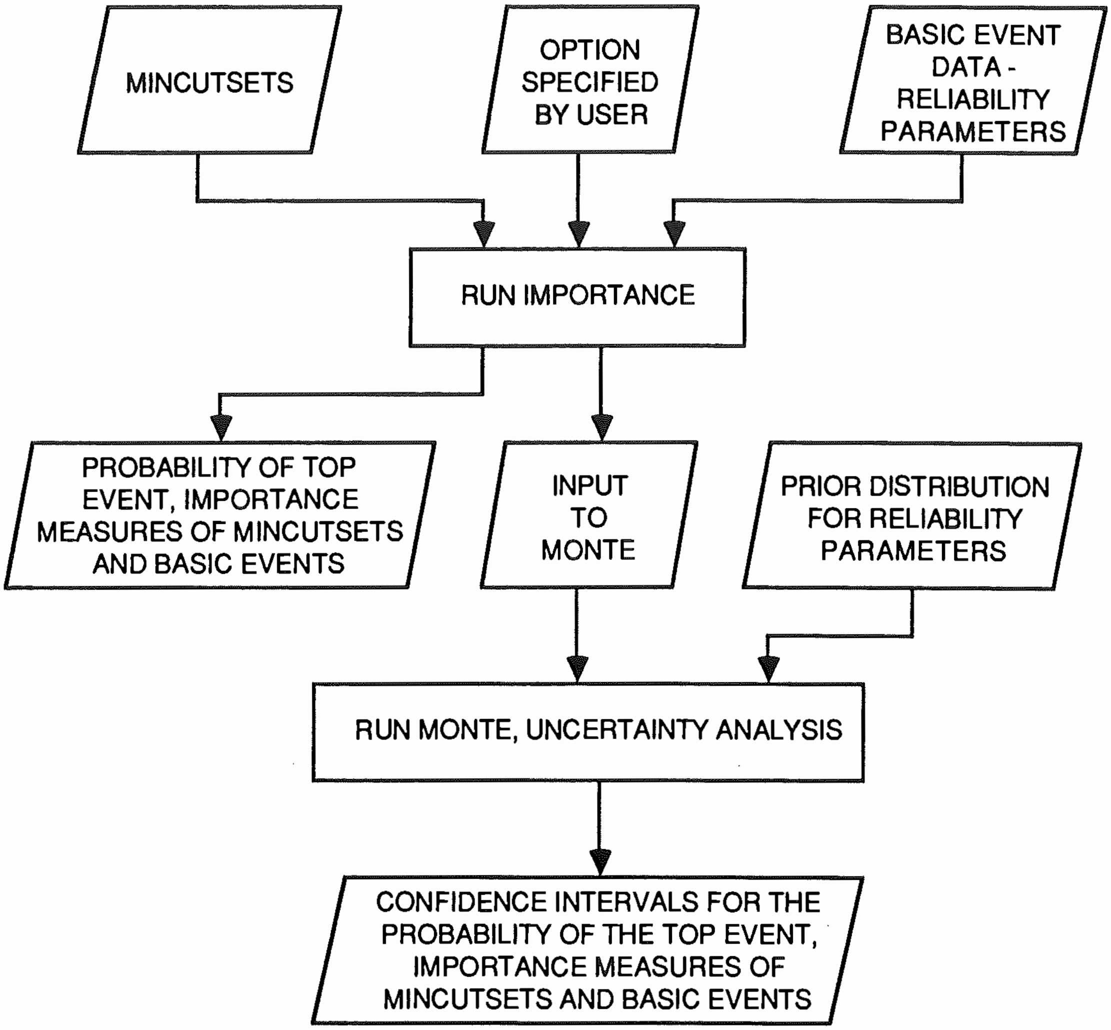

MINCUTSETS
OPTION SPECIFIED BY USER
BASIC EVENT DATA - RELIABILITY PARAMETERS
RUN IMPORTANCE
PROBABILITY OF TOP EVENT, IMPORTANCE MEASURES OF MINCUTSETS AND BASIC EVENTS
INPUT TO MONTE
PRIOR DISTRIBUTION FOR RELIABILITY PARAMETERS
RUN MONTE, UNCERTAINTY ANALYSIS
CONFIDENCE INTERVALS FOR THE PROBABILITY OF THE TOP EVENT, IMPORTANCE MEASURES OF MINCUTSETS AND BASIC EVENTS

FIGURE 4-24. INPUT – OUTPUT, IMPORTANCE AND MONTE

FTAP (Willie, 1978) accepts as input a Boolean equation for each gate event in the fault tree or logic model. FTAP can accept complemented events and k-out-of-n gates. FTAP carefully checks the input fault tree for logic errors. The code can eliminate mincutsets according to probability (as well as order) which is particularly helpful in the analysis of complex systems.

The postprocessor is especially useful for analysis of reactor accident sequences produced by the event tree-fault tree approach. Features of the postprocessor include:

- conducting common cause analysis
- dropping complemented events and performing the subsequent Boolean minimization
- generating block files (i.e., sets of Boolean equations) for subsystems.

These codes have been used extensively for nuclear power plant probabilistic risk analyses. Applications include:

- event tree-fault tree analyses
- seismic PRA analysis
- AC power reliability studies
- heavy-load drop analyses, and
- prioritizing system design changes and upgrades with cost constraints.

IMPORTANCE (Dunglinson and Lambert, 1983) accepts as input:

- basic event data, i.e., failure rates, repair times, and inspection intervals
- mincutsets from FTAP

and generates as output:

- Top Event occurrence frequency
- mean time to occurrence of the Top Event
- system unavailability, and
- the ranking of basic events and mincutsets according to various importance measures.

Dunglinson and Lambert (1983) show how rankings of basic events and mincutsets can suggest system design changes and/or procedural changes to improve safety and/or reliability.

IMPORTANCE has been revised so its calculations can perform reliability analyses of control systems. The code can distinguish between two types of basic events:

- initiating events and
- enabling events.

As has been discussed, initiating events are sources of system disturbances and place demands on the control system to respond. Examples include:

- control devices or sensors failing high, low, or reversed, and
- loss of utilities (e.g., instrument air, cooling water, or electricity.)

Enabling events represent failures of system mitigative actions, either active or passive, which permit initiating events to cause system failure.

Examples include:

- control device inactive
- pressure relief valve jammed closed, and
- interlock relay fails to open.

Enabling events do not cause disturbances but inactivate protective features.

By conducting a Monte Carlo analysis, MONTE performs an analysis of Top Event uncertainty due to uncertainties in the basic event data. Two probability distributions are allowed for the basic events' uncertainties:

- lognormal and
- normal.

Confidence intervals are generated for the Top Event probability and for basic
event and mincutset importance rankings.

Special references for this section are:

- *Nuclear Safety Analysis Computer Program Description Document*, Boeing Aerospace Company, Report D2-118655-1, February, 1979.

- Willie, R., *Computer Aided Fault Tree Analysis*, Report No. 78-14, Operations Research Center, University of California, Berkeley; Report No. UCRL-13981, Lawrence Livermore National Laboratory, 1978.

- Dunglinson, C. and Lambert, H.E., "Interval Reliability for Initiating and Enabling Events," *IEEE Transactions on Reliability*, Vol. R-32, June 1984.

For information concerning the AutoCAD/TREEGEN computer code package, contact

Tom Morgan
ERIN Engineering & Research
1850 Mount Diablo Blvd., Suite 600
Walnut Creek, CA 94596
(415) 944-7077

CHAPTER 5
EFFECTS, CONSEQUENCES, AND LOSSES MODELING

The ultimate determination of the expected (or, often, worst-case) losses or loss probabilities resulting from a mishap requires the modeling of several events: (1) Generally, a structure, container, or other equipment fails and energy and/or material is released; (2) The energy or material propagates to exposed people, systems, facilities, or the environment. If the material is flammable or explosive, it may be ignited immediately, or it may find an ignition source at some time and distance from its origin; (3) Possible levels of harm then accrue to the exposed people and property, and losses accumulate, due to immediate forces, overpressures, burning, toxicity, etc., or to delayed effects of toxicity, carcinogenicity, and so on. The models required for each of these events are necessarily specialized to the events and only general remarks can be made here about their characteristics. The development of these models for NASA systems can be the most complex aspect of the risk assessments of these systems, and can require physical and chemical, as well as engineering, analyses and simulations, and perhaps also tests, for their construction and input data development.

5.1 STRUCTURE, CONTAINER, OR EQUIPMENT FAILURES AND EFFECTS

Such failures can be due to many possible "external" causes, such as an accident (e.g., a vehicle crash), a fire in another container, or an earthquake; or to "internal" causes, such as an undetected structural defect in the container or the vehicle or mishandling in its use or maintenance.

For a container, for example, the analysis of such failures frequently involves comparing the impinging loads developed in the postulated mishap with the strength of the container. For most external causes of mishaps a dynamic situation is involved and the loads tend to be impact-induced. Examples are overpressures of explosions or the collision of one vehicle with another vehicle or a fixed object, leading to rupture of the container due to direct impact or overturning, or from penetration. These and other effect scenarios are readily treated by analysis. Estimates can also be made of the size of the opening in the breached container as a result of the explosion or impact, and then of the resulting rate and quantity of material release.

5.2 ENERGY OR MATERIAL PROPAGATION

The propagating energy or material can lead to various consequences. Fragments of various sizes and velocities can be ejected over large distances and directions with possible important dependencies on wind conditions. Volatile materials dispersed in air can cover areas orders of magnitude larger than when they were contained. A material in this state may be flammable, explosive, toxic, suffocating, corrosive, or carcinogenic.

In the event of a release of a liquefied gas or volatile liquid, for example, the escaping material will spread, evaporate, mix with the air surrounding the spill, form a cloud, and move downwind. The details of the spreading and cloud formation depend on the rate of release of the material; its density, vaporization rate and buoyancy; and on meteorological and terrain conditions. The cloud that is formed is characterized by its size and concentration at any location relative to the release point and at any time after release.

A number of mathematical models have been developed that attempt to describe these complex events. The models differ significantly from one another in sophistication, because of their approximations and assumptions in their characterizations of the source (point or area source, instantaneous or continuous release) or of the manner of spreading and air entrainment. Also, since input data on material properties are lacking for the majority of materials, data for "similar" materials are often used, giving rise to errors of uncertain magnitude in the predictions of the materials' behavior.

In order for a cloud of material in its vapor phase to burn or explode, for example, its concentration must be within its flammable limits, and an ignition source must be present. A fire or explosion gives rise to thermal radiation or overpressure, and impulsive forces that can harm nearby people and property. The flammable limits of many materials are known. The explosive effect of a material is expressed in terms of energy release, e.g., TNT equivalency, and can be estimated from the heat of combustion of the material, if this property is known. Meteorological conditions, structures, terrain features, etc., can give rise to areas where there is focusing or blast enhancement and also to areas where little damage occurs. Asymmetric initiation of a vapor cloud can give rise to enhanced blast in one direction. Predictions of fire and explosion effects tend to be conservative, since calculations often consider the worst case. However, it is sometimes possible to draw on past experience and testing to establish a less extreme, credible energy release case.

5.3 ACCUMULATION OF LOSSES

Given the distribution among the exposed people and property of the possible levels of thermal energy, overpressure, toxic concentration, etc., resulting from a possible mishap, it is then necessary to determine the expected losses or loss probabilities in the exposures.

For toxic materials, the effects of various concentrations on people and other biota are known for only a fraction of such materials. Moreover, much of this information has been developed for occupational exposures, i.e., for people exposed on a continuous, eight hours per day basis. How large a concentration is tolerable for a single exposure resulting from an accident is known only for very few materials. Worst-case assumptions again usually must be made, often based on such understanding as may exist for wider classes of similar materials. Complicating the loss estimation problem further is whether a material's biological effect can be considered to have a "threshold" dosage, below which it can be assumed no damage will result, or whether instead any dosage at all must be assessed as potentially harmful.

Harmful dosages and dose-rates of thermal energy from fire are relatively well established, as are harmful blast overpressures and dynamic pressures from explosions, at least at the higher levels of most concern. However, the effects of lower levels of blast pressures, which, for example, can cause long-range window breakage and secondary harm to people from flying or falling glass, are not so well understood and continue to be investigated both analytically and experimentally.

Nevertheless, to conclude a risk analysis with the requisite loss estimates, the spatial and temporal distribution of the exposed people and property are defined, and the spatial, temporal, and density distributions of the impinging forces, thermal energy, toxic dosages, etc., are superposed. The probability is computed that each individual or property unit will receive at least a threshold level of effect to be considered harmful. Then, these probabilities are accumulated to lead to overall risk measures, such as the probability that each individual will be harmed; the expected number of individuals that will be harmed; or (a "risk profile") the probability that each possible number of individuals, or more, will be harmed; due to the possible occurrence of a mishap in the hazardous activity of concern.

CHAPTER 6
DATA DEVELOPMENT

In this chapter procedures are outlined for the development of the basic event data input in the quantitative analysis of fault trees. Point and interval estimates of failure rates and failure probabilities of hardware components are considered,* and the development of these data by classical or Bayesian statistical methods described. A discussion is also provided of the development of human error data. Simplified procedures for developing application or system-level failure rate and probability estimates from the basic event data are presented. These procedures are applicable to cases in which certain approximations are applicable. Finally, the special problems of establishing data on non-independent failures, in particular, common-cause failures, are noted. Major elements of the discussion reflect portions of the PRA Procedures Guide (Nuclear Regulatory Commission, 1983). Additional information and references on extended techniques are provided from sundry sources as noted.

6.1 THE GENERAL DATA DEVELOPMENT PROCESS

The fundamental problem of acquiring the data requisite to the generation of basic event rates and probabilities is addressed in this section. The presentation is in part an adaptation of material from Fragola et al. (1987). The statistical analyses of the several pertinent types of data to arrive at these rates or probabilities, together with their uncertainty distributions, are discussed in subsequent sections. The steps involved in the general treatment of the data development problem are indicated in Figure 6-1.

Basic event data fall into two main categories:** application-specific data on components of concern which derive from test and operational experience with the particular application and components under analysis, and generic data on components which derive from experience with such components or similar components in other applications. Engineering

* For some applications, in which component repairs are possible, component unavailabilities that take extended downtimes into account may also need to be considered (recall Section 4.9).

** External event (e.g., earthquakes, winds, etc., of given magnitudes) probability data development requires special analysis techniques not considered further here.

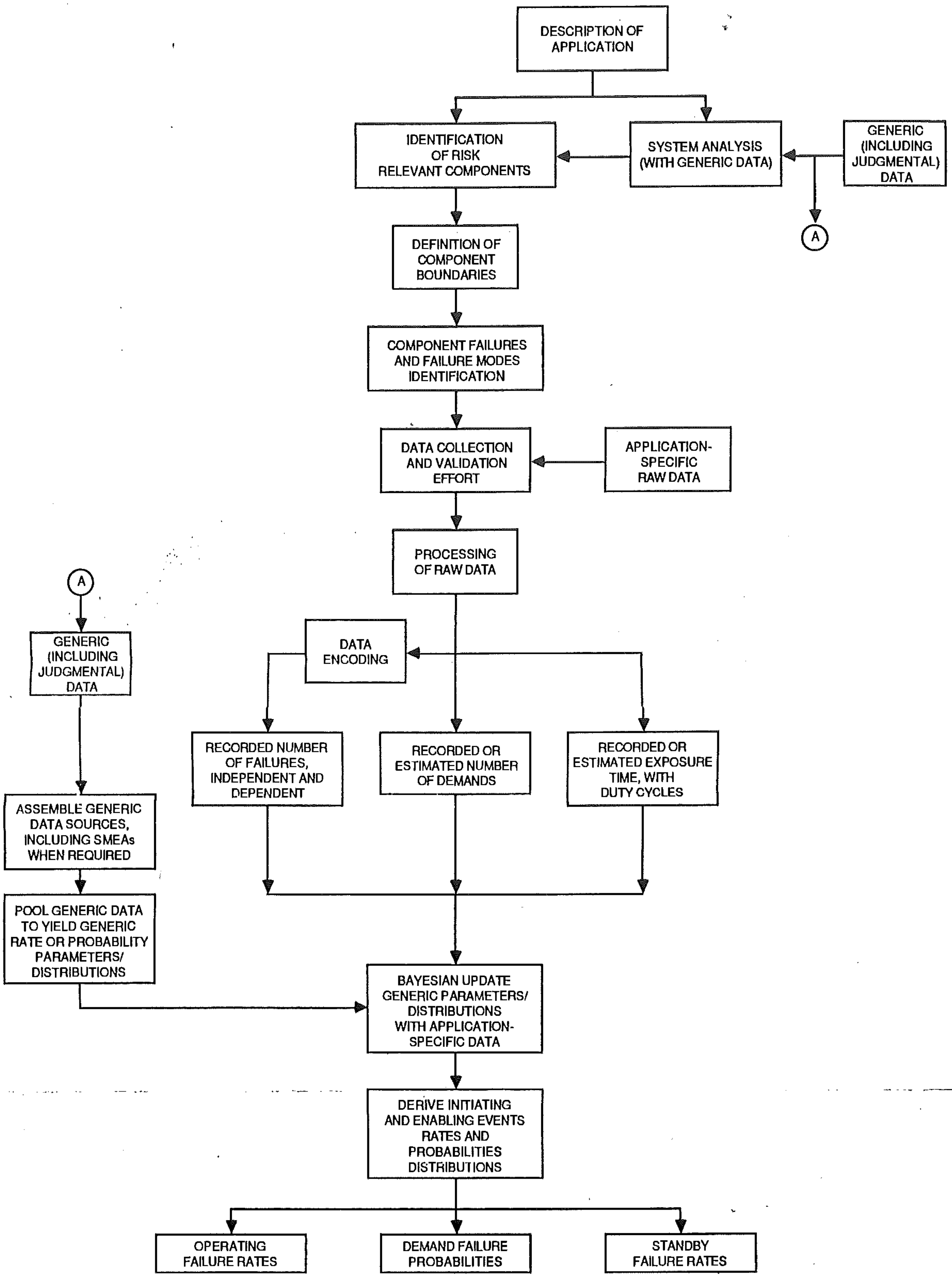

DESCRIPTION OF APPLICATION
IDENTIFICATION OF RISK RELEVANT COMPONENTS
SYSTEM ANALYSIS (WITH GENERIC DATA)
GENERIC (INCLUDING JUDGMENTAL) DATA
A
DEFINITION OF COMPONENT BOUNDARIES
COMPONENT FAILURES AND FAILURE MODES IDENTIFICATION
DATA COLLECTION AND VALIDATION EFFORT
APPLICATION-SPECIFIC RAW DATA
PROCESSING OF RAW DATA
A
GENERIC (INCLUDING JUDGMENTAL) DATA
DATA ENCODING
RECORDED NUMBER OF FAILURES, INDEPENDENT AND DEPENDENT
RECORDED OR ESTIMATED NUMBER OF DEMANDS
RECORDED OR ESTIMATED EXPOSURE TIME, WITH DUTY CYCLES
ASSEMBLE GENERIC DATA SOURCES, INCLUDING SMEAs WHEN REQUIRED
POOL GENERIC DATA TO YIELD GENERIC RATE OR PROBABILITY PARAMETERS/ DISTRIBUTIONS
BAYESIAN UPDATE GENERIC PARAMETERS/ DISTRIBUTIONS WITH APPLICATION-SPECIFIC DATA
DERIVE INITIATING AND ENABLING EVENTS RATES AND PROBABILITIES DISTRIBUTIONS
OPERATING FAILURE RATES
DEMAND FAILURE PROBABILITIES
STANDBY FAILURE RATES

FIGURE 6-1. GENERAL BASIC EVENT DATA DEVELOPMENT PROCESS
(ADAPTED FROM FRAGOLA ET AL., 1987)

judgment may be applied in the development of generic data. This is common in the establishment of the associations of the characteristics of the other applications with those of the application of concern. Engineering judgment may also provide an independent basis for the development of failure rate or probability data for some components. Various techniques have been established to support such "subjective" developments. Some are outlined later in this section.

A second categorization of basic event data reflects whether the event is a time-related failure or a demand failure. In the first case, applicable mainly to continuously operating components, the data are defined in terms of failure rates per unit time. In the second case, applicable to components that operate discretely, when required, the data are defined as probabilities of failure per operation. For the latter components, if off-operation failures are possible, perhaps also with repairs, data may also be required to estimate the probability the component will be unavailable for operation when needed (at which time it may or may not then have a demand failure).

A third categorization of basic event data is that of data on initiating versus enabling events, as was discussed earlier in Section 4. The data requirements may differ between these two types of events in that enabling events may need to endure for some time until an initiating event occurs in order to lead to an application or system-level failure (or Top Event). The parameters of the unavailability functions of components whose failures can provide enabling events thus may have to be estimated, taking into account the possible times until needed and, if repairs can be made, the possible times between repair actions and, if significant, the times to carry out repairs.

The major steps indicated in Figure 6-1 are next discussed.

6.1.1 Identification of Risk-Relevant Components

The determination of the scope of the data to be utilized, the initial step in the data development process, includes the delineation of application-specific initiating events ("initiators") that could lead to significant application or system failures. To accomplish this, systems analysts review the history of the application or system under analysis to extract the conditions that could lead to such failures, or Top Events. This is supplemented by a prognosis of events which, while having no historical basis, could have such severe consequences that they must be included in the list of potential initiators. When this identification process has been completed, the systems analyst uses his knowledge of system functional requirements to aggregate initiators according to categories of similar expected system responses. These aggregates, sometimes referred to as initiator "bins," are presented to the data analyst who reviews the events included in each bin, and establishes a frequency of occurrence for each. The associated event frequencies in a bin are then summed so as to produce an estimate of the frequency of occurrence of the previously established category (i.e., a specific grouping of initiators established by expected similar system responses). This list of initiating event frequencies is used to establish the proper occurrence rate category to be assigned to each event sequence to be considered.

After the scope of the significant initiators is determined, the systems analyst and data analyst interface in order to define the scope in areas such as specific component types, sequences, systems, subsystems, and failure modes that can give rise to the initiators and for which component failure basic events and unavailabilities will need to be developed. Generally, the systems analysis, leading to the fault tree and/or event tree model, and the data analysis are performed in parallel. The entire set of components, boundaries, and failure modes may not be known at this stage of the analysis. Generic data then are used to define a basic set of component types and modes of failure. The data analysts can then proceed with the data collection and begin sorting the data into the generic data-indicated component sets, knowing that this will constitute the minimum necessary data base.

6.1.2 Definition of Component Boundaries

Components in a major system or other application can be comprised of a number of subcomponents or piece parts, can be considered as part of a larger "component" defined by a function, or sometimes can be separated into passive and active function categories. These options for defining what should constitute a particular component must be considered by the systems analyst in the modeling of the system elements. The determination of where one component ends and another begins constitutes the definition of a component "boundary."

The boundary decisions are of great importance in the data analysis since they provide guidance to the separation and combination of raw data into component-relevant failure rates and probabilities. Without such guidance, the data analysts would have to use only their judgment as to the most likely (conventional) or logical (function-based) boundaries and build the data set accordingly. Ultimately, however, this may be inappropriate since the components modeled may not include features that the data development has assumed and the data therefore may not correctly apply to the components in the models. Furthermore, the fault tree model's definitions of component boundaries do not necessarily correspond to those considered in recording component failures in the raw data set. It is essential that the systems analyst and data analyst interface with each other to reconcile these differences and establish definitions of component boundaries that are clearly understood and acceptable to both. In some cases, the resulting definitions may make it necessary to combine basic events in the fault trees because the data cannot be obtained at the level originally conceived of by the systems analysts. Also, there may be other limiting factors directly affecting component boundary definitions. These factors are due primarily to the properties of generic data sources and are dependent upon how the sources have defined the component attributes. The boundary determinations are therefore not only made by considering the level of basic event detail, but also by considering the availability and detail of the application-specific failure (and repair) records and the previously determined generic data that will be used in the analysis where necessary.

6.1.3 Identification of Component Failures and Failure Modes

The definition of the boundaries of the components under consideration greatly assists the determination of the ways in which these components can fail to function. However, "failure" is a term which can mean different things to different people. Therefore, the types of failures should be defined and classified to prevent confusion. Fundamentally, the distinctions are made on the basis of failure severity, or the degree to which the component has been disabled. For example, the following severity categories may be employed:

Catastrophic — The component is completely unable to perform its function.

Degraded — The component operates at less than its normal performance level.

Incipient — The component performs within its design envelope but exhibits characteristics that, if left unattended, will possibly develop into a degraded or catastrophic failure.

Such failure severity categories provide criteria for the definition of component failure modes, that describe the basic types of failures a component can experience, such as "fails to open" or "fails to run."

The failure modes and component boundaries specified by the systems analyst in the determination of the scope is a major factor to be considered by the data analyst during the review of the raw data in the initial application-specific data reduction task. Failures and unavailabilities should accurately reflect the component's performance as portrayed in the basic event. Understanding the failure process is necessary in determining a cause of failure and thus the particular component that failed.

6.1.4 Categories of Failure and Unavailability Data Calculations

As has been noted, failure modes can be organized generally by whether the failure is time-related or demand-related. A time-related failure rate is based upon those failures that occur as a result of time elapsing, such as a failure to continue to function or so as to give rise to a spurious operation. Such failure rates can be calculated for components that either are normally in an operating state or are most often in standby.

When a component in standby is called upon to function and fails to perform, the failure is demand-related and the number of such failures out of the total requests to operate (demands) is called the demand failure probability. Examples of failure modes that correspond to demand-related failures are "motor failure to start" and "switch failure to open on demand."

Failure rates and probabilities must be constructed from raw data if an application-specific data base is used. These same parameters, however, are often already available "pre-calculated" from generic data sources and the work in this case is focused on correctly applying the ready-made data to the situation at hand.

Apart from the rate or probability with which a failure will occur in a given component, the analyst must also define the amount of time a component is not available to perform its function, if its inability to operate is an enabling event. Unavailability of a component in some systems may not only occur due to a failure, but also due to the component's undergoing testing and maintenance. Application-specific outage data are the best source for such unavailability information, which is usually expressed as the ratio of the time a component is out of service due to test and maintenance to the total time in the period under consideration (with statistical uncertainties also taken into account). Application-specific outage data can be gathered in a number of ways. The method selected may be dependent upon the resources available to the analysis, both the amount and nature of data and the number of people and funding for the analysis.

6.1.5 Collection and Processing of Raw Data

The main motivation for the selection of raw application-specific data is the recognition that actual records of experience provide the best source for the computation of failure rates, failure probabilities, unavailabilities, and other reliability and, perhaps, maintenance factors required for the analysis. The information extracted from the records must be carefully documented to provide traceability back to the raw data from the calculated results.

An initial sort of the data is performed, usually by subsystem or other separable application area, to structure the data. Then each record is reviewed, first to determine whether it actually describes a failure, then to separate "relevant" analysis records (failure- or maintenance-related; association with a subsystem pertinent to the analysis; etc.) from "irrelevant" records (not related to failure or maintenance; associated with a subsystem not considered pertinent). Relevant records are then scrutinized to characterize each failure or maintenance action of interest that occurred.

6.1.6 Data Encoding

The minimum set of required elements to be extracted from the raw records to allow for constructing the failure and unavailability data sets are: the time of failure occurrence, subsystem or application area identification, component type, component identification, and a narrative description of the failure and any repair. The time of occurrence may be an important element because it indicates whether or not the component failure occurred during the time period specified for the analysis, and, when appropriate, the phase of system operation in which it occurred. The subsystem identification and component type are required inputs for the model, while the component identification provides traceability back to the raw data. When the component identification and the time of occurrence are utilized together, the identification of any duplications of records is easily accomplished. The narrative description is reviewed and analyzed to determine factors such as failure modes, failure severities, and whether or not a relevant failure has occurred. The narrative section usually provides the only means by which the data analyst can determine if the documented event is a true failure, a routine maintenance action, or a specified test. In addition, the narrative description enables the determination of the type of failure, essential to the segregation of demand-related failures from time-related failures. Failures for each particular component type in each subsystem are then characterized by whether the failure is time - or demand - related (as determined through the failure mode) and next by the severity classes defined earlier; e.g., Catastrophic, Degraded, and Incipient. The Catastrophic failures are then counted to provide numerators for failure rates (time) and probabilities (demand).

6.1.7 Demand Spectrum and Exposure Time Determination

When a number of failures has been determined for a specific component type in a particular subsystem or application area, the next step is to estimate for each component type the number of demands and/or the total operating time during which the failures occurred. A method for estimating these parameters is required unless the operating times or demands for a component have been recorded during tests and operations. Otherwise, system information, including operating characteristics and design configurations, test information, and operating history are required to produce the exposure estimates.

The information for each component type that is obtained from operating history includes the duty cycle and/or the total number of operations and the total time in each mode of operation during the time period.

Testing information provides data on demands and operating hours during tests. When applying the test information in estimates of the number of demands and the operating time, the data analyst must possess or have access to enough operational knowledge that a truly representative number of demands and total operating hours can be estimated.

6.1.8 Generic Data

The compilation of generic data requires data source review, data comparison, and data selection. Subjectively-derived engineering estimates as well as objective experience data from similar applications of the components of concern may be considered.

6.1.8.1 Objective Generic Data

Using the component listing as a guideline for the type of information that needs to be compiled, the data source review requires that the analyst obtain the generic sources and investigate them thoroughly. The major difficulties in accessing generic data sources arise from the time-consuming nature of the task and the comprehensiveness of the search for sources required to ensure that all applicable sources have been addressed. Many of the generic data sources will be published documents that are readily available (see, e.g., Table 6-1). Sources more specifically related to NASA systems may also be found; e.g., Department of Defense missile and space programs. The proper handling of classified data may sometimes be required. As these and other generic sources are reviewed for pertinent data, the key points that have been emphasized above should be kept in mind:

- The component boundary
- The definition of the failure modes (e.g., what failure modes are included and excluded)
- Time-related vs. demand-related rates
- The confidence interval expressing uncertainty, and
- The data source(s) behind the data.

As has been noted, the component boundary defines the equipment that is considered as part of the component type for which a failure rate has been designated in the source. The failure modes defined for a component for the analysis being conducted may differ from those cited in the data sources. For example, it is important to know whether to equate a "failure to open" with a "failure to operate" mode, and the latter may include other modes, such as "plugged" or "blocked," besides the failure mode of interest. The failure rate is usually considered to be the time-related rate of failure, but for some modes, the failure probability is expressed in terms of the probability of failure on demand. The time units must of course be understood explicitly. Different sources may also provide data error bounds in different manners. For example, a 90% confidence interval may be given in one source, an 80% interval in a second, a variance or standard deviation in a third, and an error factor in a fourth. Comparisons of the relative confidence of alternative data values depends on understanding any such differences.

- *IEEE Guide to the Collection and Presentation of Electrical, Electronic, Sensing Component and Mechanical Equipment Reliability Data for Nuclear Power Generating Stations*, IEEE Std 500 - 1984, The Institute of Electrical & Electronic Engineers, Inc., December 1983.

- Arno, Robert G., *Nonelectronic Parts Reliability Data*, NPRD - 2, ITT Research Institute, 1981.

- *Reactor Safety Study - An Assessment of Accident Risks in U.S. Commercial Nuclear Power Plants*, WASH - 1400 (NUREG/75/014), U.S. Regulatory Commission, Washington, DC, October 1975.

- *Reliability Prediction of Electronic Equipment, Military Handbook - 217E,* Department of Defense, Washington, DC 20301, January 1982.

- Generic databases of the various nuclear plant PRAs.

- For transportation accident data, DOT databases.

Table 6-1. Representative Generic Component Data Sources

Finally, some generic data handbooks use a combination of other data sources in deriving their failure rates. This must be recognized so that the same original source is not doubly counted in generating an aggregate failure rate.

As the failure modes, failure rates, error bounds, and other factors are extracted from the data sources, it is helpful to organize the data to facilitate its later use. Figure 6-2 shows an example of the type of format that has been used in recording the extracted generic data and, at the same time, correlating it to components and failure modes, in some nuclear plant analyses. To establish a one-to-one correspondence between the failure modes of concern and those found in the sources, it is useful to construct a failure mode hierarchy that clarifies the degree of correspondence. Figure 6-2 also gives an example of such a hierarchy in the second and third columns of the matrix.

COMPONENT TYPE	TAXONOMY FAILURE MODE	SOURCE FAILURE MODE	FOREIGN DATA SOURCES			U. S. DATA SOURCES						
			BIBLIS B	EUROPEAN	OREDA	ASEP	IEEE-500 1984	NPRD-2 (RAC)	PSA UPDATE	WASH-1400	SEABROOK PRA	OCONEE PRA
C: MOTOR OPERATED VALVE	C.1 ALL MODES	C.1 ALL MODES			8.1×10^{-4} /hr*		$.13 \times 10^{-7}$ /hr					
	C.1a FAILURE	C.1a FAILURE		3.9×10^{-7} /hr								
	C.1.1 CATASTROPHIC	C.1.1 CATASTROPHIC			3.5×10^{-6} /hr		$.13 \times 10^{-1}$ /hr					
	C.1.1.1 FAILURE TO OPERATE	C.1.1.1 FAILURE TO OPERATE				3×10^{-3} /d	6×10^{-3} /d			1×10^{-3} /d	4.3×10^{-3} /d	6.4×10^{-3} /d
		C.1.1.2 FAILURE TO REMAIN OPEN				1×10^{-4} /d			2×10^{-7} /hr	1×10^{-4} /d		
		C.1.1.3 FAILURE TO CLOSE							1×10^{-5} /hr			
		C.1.1.3a FAILURE TO CLOSE ON DEMAND WHILE INDICATING CLOSED									1.07×10^{-4} /d	
	C.1.1.4 SPURIOUS OPERATION				* BASED ON VERY SMALL SAMPLE							

FIGURE 6-2. FAILURE MODE AND RATE CORRELATION MATRIX
(ADAPTED FROM FRAGOLA ET AL., 1987)

6.1.8.2 Subjective Development of Generic Data from Subject Matter Experts

Subjective assessments by Subject Matter Experts (SMEs) of failure rates and probabilities, maintenance times, etc., are often employed to augment a generic data set. While the assessments are focused on a specific application, their use in risk analysis is in the development of prior distributions that will be revised with any available objective data. Thus, they play the same role as generic data and so are included among them.

The structured elicitation of subjective probability data from SMEs to supplement objective generic data where necessary is a common practice in nuclear power plant risk assessments and other reliability and risk analyses. Several techniques for conducting such elicitations are described in Section 6.3.3.6 below. That section and also Fragola et al. (1987) may be referred to for a discussion of some of the main problems that may arise, including the problem of bias and the tendency of SMEs to be overconfident in the precision of their judgments.

6.1.8.3 Generic Data Integration

The completion of the data extraction and documentation allows the data from the various sources to be compared to select the single source or combination of sources whose data will best meet the needs of the analysis. Analytical methods for combining the data from several sources are described in Section 6.3.3.5, below. The factors to be considered in this comparison process are the following:

- The ultimate data source(s)
- The confidence interval
- The system or application that data come from
- The failure mode correspondence
- The component utilization and boundary, and
- Time- or demand-related failure modes.

There is no strict formula for the application of these, and potentially other, factors in the evaluation of generic source data. Rather, a tradeoff among these factors must generally be made to select the most applicable data. The selection process should be carefully documented to preserve the selection decisions and the logic behind them.

6.2 CLASSICAL STATISTICAL INFERENCES OF COMPONENT FAILURE PROBABILITIES AND FAILURE RATES FROM TEST AND EXPERIENCE DATA

When a sample of test or operational experience data is available it may be possible to establish satisfactory estimates of a primary event's parameter values by means of classical statistical inference techniques. Generally, not merely an estimate of a parameter's nominal value is desired, but also the distribution of the uncertainty in the estimate due to the randomness of the sample. A number of cases arise depending on the character of the event.

6.2.1 Failure of a Component That Meets Discrete Demands

The probability of failure per demand, p, of a one-shot or multiple-shot component (such as a pyrotechnic device or a relay) that performs a discrete operation on demand is estimated as

$$\hat{p} = r / n \qquad (6\text{-}1)$$

where n is the number of demands made during test and operation of the component and r is the number of failures observed (without replacing any failed units) in meeting these demands.*

Upper and lower approximate 100(1- α)% confidence limits defining interval estimates for p when r is small (even zero) and n is large are*

$$p_U(a) = X^2\ (2r+2;\ 1\text{-}\alpha/2) / 2n \qquad (6\text{-}2)$$
$$p_L(a) = X^2\ (2r;\ \alpha/2) / 2n$$

* It is to be noted that particular care may be necessary if the number of demands n over some period of time in which the number of failures r was accrued must be estimated, rather than directly counted. The component's "duty cycle" is involved. This is an example of the not uncommon "exposure" estimation problem in statistical risk inferences.

* Type I censored data is assumed, with the data taken for a selected number of <u>operations</u> or <u>demands</u>, n. If Type II censored data are obtained, with termination of the data acquisition at a selected number of <u>failures</u>, a somewhat tighter approximate upper confidence bound is applicable:

$$p_U(\alpha) = X^2\ (2r;\ 1\text{-}\alpha/2) / 2n \qquad (6\text{-}3)$$

where X^2 (d;α) is the tabulated 100 α% percentile of a chi-square variate with d degrees of freedom.

Exact values for a set of percentiles for p for smaller n (up to 100) are tabulated in Huebel and Myers (1976). (See also Pratt, 1968, I and II, and Blyth, 1986, for normal approximations.)

A standard assumption for a complete distribution for the uncertainty in p (assumed to be represented by the random variation in the sample estimate) is that of the log-normal.** An alternative sometimes assumed is the beta distribution (see, for instance, Mann, Schafer, and Singpurwalla, 1974). A complete distribution for each component failure probability's uncertainty is needed as a basis for an analytical (when feasible) or Monte Carlo sampling development of system-level interval estimates of Top Event probabilities.

If a sufficient data sample size is available, it is possible to check the reasonableness of the log-normal (or other) distributional assumption with a test of goodness-of-fit (see, e.g., Winkler and Hays, 1975).*** The X^2 (chi-square) test is the most used, but requires a large sample size. The Kolmogorov-Smirnov test can be applied to smaller samples. It is not to be expected that many NASA tests or operations will generate sufficient data to enable either of these tests, but they are described here nevertheless for use in such cases as will arise.

6.2.1.1 X^2 Test

Break the total sample into at least 25 subsamples, say, and obtain an estimate $\hat{p}$ for each subsample. Consider the range of values of $\hat{p}$ and decompose the range into at least 5 subranges or intervals I_i each containing n_i values of $\hat{p}$, where n_i is at least 5. Let k be the number of such intervals and let the total sample size be

** The use of an "uncertainty distribution" is at best controversial with classical or "frequentist" statisticians. Bayesians employ it with freedom, as will be seen later.

*** Since all values of p must be between 0 and 1, the tail of any fitted log-normal distribution, which theoretically extends to infinity, must of course become negligible before p values near 1 are reached. Since the domain of the beta distribution is a finite interval, this is not of concern if a beta model for the uncertainty in p is assumed.

$$N = \sum_{i=1}^{k} n_i \qquad (6\text{-}4)$$

From the assumed log-normal uncertainty distribution compute for each I_i the probability Π_i that the true value of p falls in I_i. Multiply Π_i by n to get the number e_i of $\hat{p}$ values to be expected to fall in I_i if the hypothesized distribution is correct.

Now compute the sum S over the I_i of $(n_i - e_i)^2 / e_i$, which is approximately X^2-distributed. The hypothesized distribution is rejected as incorrect with confidence 100 (1 - α) % if S is larger than the tabulated value of the 100 (1 - α)% percentile of X^2 with k-1 degrees of freedom. If S is smaller than the 100 (1 - α)% percentile value, then the hypothesis can be accepted with confidence 100 α %. Clearly, S should be smaller than X^2 percentile values corresponding to relatively large α for reasonably high confidence in the goodness-of-fit of the hypothesized distribution to the data.

6.2.1.2 Kolmogorov-Smirnov Test

This test can employ a smaller sample size because grouping of the $\hat{p}$ values into intervals is not required. Let $\hat{p}_j$, j = 1,2,...,J, be these values ordered by increasing magnitude. A cumulative distribution function for the uncertainty due to sample value randomness is then the step-function

$$\hat{F}(p) = j / J , \qquad \hat{p}_j \leq p \leq \hat{p}_{j+1} \qquad (6\text{-}5)$$

with $\hat{p}_0 = 0$, $\hat{p}_{J+1} = +\infty$, by definition.

The cumulative distribution function F(p) corresponding to the hypothesized uncertainty distribution is then compared to $\hat{F}(p)$ by considering the maximum difference D between corresponding F and $\hat{F}$ values. The hypothesized distribution is rejected as incorrect with confidence 100 (1- α)% if D is greater than a tabulated value $D^{(J)}(1 - \alpha)$ (which is independent of any particular function F). If D is smaller than $D^{(J)}(1 - \alpha)$ then the hypothesized distribution is accepted with confidence 100 α %.

6.2.2 Failure of a Component that Operates Continuously

The time-to-failure t of a continuously operating component is usually assumed to be distributed as a one-parameter exponential, $\lambda e^{-\lambda t}$, with a constant failure rate, λ , or, more generally, when a power function variation in the rate over time is modeled, as a two-parameter Weibull distribution,

$$f(t;\, \beta, \lambda) = (\,\beta\, \lambda^{\beta}\, t^{\beta-1}\,)\;\; e^{-(\lambda t)^{\beta}} \tag{6-6}$$

(see, e.g., Mann, Schafer, and Singpurwalla, 1974). In the exponential case (which is also Weibull with $\beta = 1$), the probability of failure by time t, i.e., the unreliability of the component, is $1 - e^{-\lambda t}$. In the Weibull case, it is $1 - e^{-(\lambda t)^{\beta}}$. The Weibull generalizes the exponential by reflecting an increasing ($\beta > 1$) or decreasing ($\beta < 1$) failure rate, $\lambda^{\beta}\, t^{\,\beta-1}$.

Exponential Case

The classical point and approximate confidence bound estimates for λ from life test or operational experience data are the same as for the binomial parameter p with the number of observations n replaced by the total test or operating time t:

$$\hat{\lambda} = r\,/\,t \tag{6-7}$$

$$\begin{aligned} \lambda_U\,(\alpha) &= x^2\,(2r + 2;\, 1 - \alpha/2)\,/\,2t \\ \lambda_L\,(\alpha) &= x^2\,(2r;\, \alpha/2)\,/\,2t \end{aligned} \tag{6-8}$$

As was noted for the estimate for p, the upper bound is for Type I censored data (observation until a selected time t). For Type II data (observation until a selected number of failures), 2r+2 is replaced by 2r as the degrees of freedom of the X^2 variate. Note that the Type I case must be considered if a failure rate bound is desired before any failures occur (r=0).

More accurate (but more complex) expressions for the bounds on λ in the two censoring cases have been developed (see, e.g., Wright, Engelhardt, and Bain, 1978, for the Type I censored data case) which may be of special value when the data sample is small. Also, bounds have been derived for mixed censoring (Fairbanks, Madsen, and Dykstra, 1982). These bounds may be useful when, for instance, Type II test data are available for combination with inherently Type I experience data. Other variations that have been treated are the case in which failed items are replaced or repaired during the test or operations, and the case in which an additional "threshold" parameter is introduced whose effect is to model a zero probability of failure (e.g., due to fatigue stress) until some particular time has elapsed (Wright, Engelhardt, and Bain, 1978).

A log-normal is commonly assumed as a model for λ 's uncertainty distribution just as it is for p's. Note that whereas it was essential that the tail of p's distribution became negligible for p values approaching 1 (since $p \leq 1$, of course) no such requirement exists for λ 's distribution.

Weibull Case

The probability of failure by time *t* is now $1\text{-}e^{-(\lambda t)^{\beta}}$. The (maximum likelihood) point estimates of β and λ from sample data on failures versus time are not obtainable as closed-form expressions, but require the numerical solution of certain equations (Bain 1978, p. 109; note that $\theta = 1/\lambda$ is considered instead of λ). A graphical technique for obtaining point estimates of the Weibull parameters from sample data is described in Mann, Schafer, and Singpurwalla (1974, pp. 214-217). See also the extensive delineation of Weibull parameter estimation techniques in Abernethy et al. (1983).

Expressions for confidence intervals for β and λ require complex calculations also. However, in the Type II censoring case (data up to a selected number of failures) somewhat simplified approximations are available (Bain, 1978, p. 264-280). These apply to any sample size (number of items under observation until a selected number of failures occurs). For larger sample sizes, when a normal approximation applies, still simpler procedures can be used (Bain, 1978, pp. 283-285). With Type I data censoring (in which data acquisition is terminated at a selected time), less complete results are available. Sirvanci and Yang (1984) discuss point estimates of β and λ and their variances but except in the large sample case (when approximate normality may be assumed) the development of confidence intervals appears yet to be accomplished.

6.3 BAYESIAN INFERENCES

The applicability of classical inference techniques depends essentially on the availability of an adequate data sample of demands and failures to meet demands, or of operating and failure times. Except in controlled testing of many identical items, or repeated testing of a smaller number of perfectly repairable items, such a sample is often hard to come by. When it is attempted to use operational experience data it is frequently a problem that the failure probabilities of the items observed vary dependently, as, for instance, corrections are made to eliminate or decrease the likelihood of occurrence of previously observed failure modes. Thus, sources of data that can at least supplement classically applicable data may be necessary if useful statistical inferences of failure probabilities or rates are to be established. Bayesian procedures provide the mechanism for developing and using such data, including any sparse test or operational data that may exist. In particular, Bayesian procedures can be applied in cases in which no failures have occurred, and even before any experience at all has accrued.

It is also to be noted that Bayesian procedures facilitate, and much better rationalize than can classical methods, the propagation of the uncertainties (or, classically, possible ranges of errors) in the components' failure rates or probabilities to the uncertainties in the system-level Top Event rates or probabilities. In Bayesian analyses, component-level uncertainty distributions are established as a matter of course and these are convolved systematically in one unambiguous way or another to arrive at system-level uncertainty distributions. Purely classical methods are relatively awkward, involving either (1) only component-level confidence intervals to work with which are not directly associated with specific distributions, or, (2), attempts to develop system-level confidence bounds directly from the observed data on the components making up the system. In the first case, component-level distributions must be assumed at least implicitly, when "fiducial" techniques are employed (Mann, Schafer, and Singpurwalla, 1974, pp. 487-490); or, as was done in the preceding section on (tainted) classical statistical inference procedures, assumed explicitly, in which case a Bayesian concept is in essence introduced. In the second case, a relatively simple model (e.g., a series model) of the system's failure probability in terms of its components' failure probabilities must apply (see, e.g., Mann, 1974; Winterbottom, 1974; Basu and El Mawazina, 1978). See also Maximus (1981) for other more or less useful approximate classical techniques of this second type that, in effect, reduce a system to an "equivalent component" with "equivalent test data" developed from the actual component test data, to which standard inference methods are then applied.

A useful reference for general Bayesian concepts and techniques as well as added details on most of the subjects presented here is Martz and Waller (1982).

The following sections treat the Bayesian development of variously applicable component-level uncertainty distributions and their propagation to system-level uncertainty distributions. The material presented incorporates much of Section 5.2.2 of Nuclear Regulatory Commission (1983), with additions on propagation methods and on techniques recently developed for making fuller use of engineering failure analyses of certain types of components, and error analyses of human actions.

6.3.1 Bayesian Concepts

The Bayesian approach is similar to the classical approach in that it yields "best" point estimates and interval estimates, the intervals representing ranges in which, one is confident, the parameter value really lies. It differs in both practical and philosophical aspects, though. The practical distinction is in the incorporation of belief and information beyond that contained in the observed data; the philosophical distinction lies in assigning a distribution that describes the analyst's belief about the value of the parameter. This is the so-called prior distribution.

The prior distribution may reflect a purely subjective notion of probability, as in the case of a Bayesian degree-of-belief distribution, or any physically caused random variability in the parameter, or some combination of both. Physically caused random variations in a parameter like a failure rate may stem from system effects, operational differences, maintenance effects, environmental differences, and the like. The distribution that describes this physically-caused random variation in the parameter is sometimes referred to as the "population variability" distribution (Apostolakis et al., 1980) and can be represented by a Bayesian prior distribution. However, such random variation in the parameter can also be modeled by classical methods, using compound distributions in which the population-variability distribution becomes the mixing distribution. On the other hand, if the prior distribution embodies subjective probability notions regarding the analyst's degree of belief about the parameter, the Bayesian method is the appropriate framework for making parameter estimates. A comparative discussion of both interpretations of the notion of probability, the subjective and the relative frequency interpretations, is given by Parry and Winter (1981).

Whether the analyst does or does not have objective relative frequency data, he will often have other information based on engineering designs, related experience in similar situations, or the subjective judgment of experienced personnel. These more or less subjective factors will also be incorporated into the prior distribution--that is, into the description of his prior knowledge (or opinions) about the parameter.

The Bayesian method takes its name from the use of Bayes' theorem and the philosophical approach embodied in the 18th-century work of the Rev. Thomas Bayes. Bayes' theorem is used to update the prior distribution with directly relevant data. Here the term "generic data" will be used to refer to parameter-related information that is nonspecific to any particular application, being an aggregation over more than one use condition. A probabilistic risk assessment (PRA) for a particular application, of course, requires not generic data but rather estimates that are specific to the application. Bayes' theorem then updates the prior distribution with specific evidence and has the effect of "specializing" the prior to the specific application. The updated, or specialized, prior is called the "posterior distribution" because it can be derived only after the specific evidence is incorporated. The prior reflects the analyst's degree of belief about the parameter before such evidence; the posterior represents the degree of belief after incorporating the evidence. Application-specific estimates are then obtained from the posterior distribution as described below.

6.3.2 Essential Elements of the Bayesian Approach

This section considers the essential elements of the Bayesian approach to data reduction. It presents a brief discussion of Bayes' theorem, the basic ideas of Bayesian point and interval estimation, and a step-by-step outline of the procedures for obtaining Bayesian estimates.

A main benefit in using the Bayesian approach to data reduction is that it provides a formal way of explicitly organizing and introducing into the analysis assumptions about prior knowledge. This knowledge may be based on past generic industry-wide data and experience, engineering judgment, expert opinion, and so forth, with varying degrees of subjectivity. The parameter estimates will then reflect this knowledge. A second benefit is that the Bayesian approach facilitates the propagation of evaluations of component-level uncertainties to evaluations of uncertainties at the system level.

6.3.2.1 Bayes' Theorem

The fundamental tool for use in updating the generic prior distribution to obtain application-specific parameter estimates is Bayes' theorem. If, for instance, the parameter of interest is a failure rate λ (number of failures per unit time), Bayes' theorem states that

$$f(\lambda \mid E) = \frac{f(\lambda)\, L(E \mid \lambda)}{\int_0^{\infty} f(\lambda)\, L(E \mid \lambda)\, d\lambda} \qquad (6\text{-}9)$$

where $f(\lambda \mid E)$ is the posterior distribution, the probability density function of λ conditional on the specific evidence E; $f(\lambda)$ is the prior distribution, the probability density function of λ based on generic or subjective information but incorporating no specific evidence E; and $L(E \mid \lambda)$ is the likelihood function, the probability distribution of the specific evidence E for a given value of λ .

If the parameter of interest is the probability of failure on demand, p, rather than a failure rate λ , then λ is simply replaced by p in Equation 6-9. However, the likelihood function will differ for the different cases, as shown below.

In certain special cases, the integral on the right-hand side of Equation 6-9 can be evaluated analytically to give a closed-form expression for the posterior distribution. The term "conjugate prior" is used to describe the prior distribution form that conveniently simplifies the integration. For example, if the likelihood function is the Poisson distribution, then the gamma family represents the conjugate prior: the posterior distribution will be expressible in closed form as another gamma distribution. In general, a closed-form integration will not be possible and numerical techniques must be used; alternatively, the continuous prior distribution can be approximated by a discrete approximation and the integral replaced by a sum.

Numerical integration or a discrete approximation is often needed when the generic data include a precise description of a prior distribution, so that the analyst lacks the flexibility to choose a mathematically tractable form for it. For example, if a log-normal prior distribution is specified for λ and the likelihood is the Poisson distribution, then the posterior distribution cannot be obtained analytically in closed form. On the other hand, if one has incomplete information, this choice can be made from the conjugate family of distributions, which yields the mathematical convenience and resultant simplicity of a closed-form expression for the posterior distribution. Sensitivity studies can then be conducted to examine the effects of this choice.

The discrete form of Bayes' theorem is

$$f(\lambda_i \mid E) = \frac{f(\lambda_i)\, L(E \mid \lambda_i)}{\sum_{i=1}^{m} f(\lambda_i)\, L(E \mid \lambda_i)} \tag{6-10}$$

where λ_i $(i = 1,2,\ldots,m)$ is a discrete set of failure-rate values. The prior and posterior distributions are approximated by the discrete functions $f(\lambda_i)$ and $f(\lambda_i \mid E)$, respectively.

The discrete form of Bayes' theorem is mathematically convenient and is sometimes used as an approximation of the continuous form given by Equation 6-9 when the denominator in Equation 6-9 cannot be evaluated in closed form. In such cases, the range of the parameter is decomposed into a set of intervals and the probability content of each interval is then associated with a single point inside the interval.

Note that the denominator of either Equation 6-9 or Equation 6-10 can be thought of simply as a normalizing factor that makes the posterior distribution integrate or sum to unity. Thus, Bayes' theorem can be interpreted as merely saying that the posterior distribution is proportional to the product of the prior distribution and the likelihood function.

6.3.2.2 Bayesian Point and Interval Estimation

The prior distribution summarizes the uncertainty in a parameter as reflected by prior judgment and/or the generic data sources on which the prior is based. Similarly, the posterior distribution summarizes the uncertainties in the application-specific value of the parameter as reflected by the combined influence of both the prior distribution and the likelihood function. In either case, it is frequently desired to obtain a point or interval estimate of the parameter.

A Bayesian point estimate is a single value that, in some precisely defined sense, best estimates or represents the unknown parameter. Two commonly used point estimates are the mean and the median (50th percentile) of the prior or the posterior distribution. The mean of a distribution is the Bayesian estimate that minimizes the average squared error of estimation (averaged over the entire population of interest), while the median is the one that minimizes the average absolute error. Thus, either the mean or the median of the prior distribution can be used as a point estimate of the unknown generic parameter; likewise, the mean or the median of the posterior distribution can be used as a point estimate of the unknown application-specific parameter. The properties of the two estimators are discussed by Martz and Waller (1982). The mean or the median are found by conventional statistical procedures: using the prior distribution, the mean of a failure rate μ_λ is given by

$$\mu_\lambda = \int_0^\infty \lambda \ f(\lambda) \, d\lambda \tag{6-11}$$

while the median is the solution to

$$F(\lambda) = \int_0^\lambda f(t) \, dt = 0.5 \tag{6-12}$$

F(λ) denoting the cumulative distribution function. Using the posterior distribution, the prior f (λ) would be replaced by the posterior f (λ I E) in Equations 6-3 and 6-4.

Now consider the problem of obtaining an interval estimate for λ, using either the prior or the posterior distribution, depending on whether one is concerned with a generic or a specific failure rate. Suppose a probability of (1 - γ) is wanted that the interval estimate in fact includes the unknown failure rate. (For example, γ = .05 for .95 probability.) One can obtain a 100(1 - γ) % two-sided Bayes probability interval estimate of λ by solving the two equations

$$\int_0^{\lambda_L} f(\lambda) \, d\lambda = \gamma / 2 \tag{6-13}$$

and

$$\int_{\lambda_u}^\infty f(\lambda) \, d\lambda = \gamma / 2 \tag{6-14}$$

for the lower end point λ_L and the upper end point λ_U. It follows immediately that $P(\lambda_L < \lambda < \lambda_U) = 1 - \gamma$.

For a Bayesian interval estimate of an application-specific failure rate, the posterior distribution $f(\lambda \mid E)$ would replace the prior distribution $f(\lambda)$ in Equations 6-13 and 6-14. The interval estimate (λ_L, λ_U) would then be such that $P(\lambda_L < \lambda < \lambda_U \mid E) = 1 - \gamma$.

Analogous results hold when the parameter of interest is a failure-on-demand probability p rather than a failure rate λ.

6.3.2.3 Step-by-Step Procedure for Bayesian Estimation

For estimating a parameter such as a component-failure rate or a failure-on-demand probability, the steps are as follows:

1. Identify the sources and forms of generic or subjective information to be used in selecting an appropriate prior distribution for the parameter.

2. Select a prior-distribution family if none has been specified as part of the generic information.

3. Choose a particular prior distribution by reducing and/or combining the generic data from step 1.

4. Plot the prior and summarize it by determining its mean, variance, and selected summary percentiles.

5. If generic estimates are required, determine them from the prior as in Section 6.3.2.2.

6. If application-specific estimates are required, then --

   a. Obtain data representing operating experience with the specific component.

   b. Identify an appropriate form for the likelihood function.

c. Use Bayes' theorem to get the posterior distribution.

d. Plot the posterior distribution on the same graph with the prior and summarize the posterior in the same manner as in step 4.

e. Compare the prior and the posterior distributions to see the effect of the specific data.

f. Obtain the desired estimates from the posterior distribution.

7. Investigate the sensitivity of the results to the prior distribution.

These steps are next discussed in detail.

6.3.3 Determining Prior Distribution

A fundamental part of a Bayesian estimation procedure is the selection and fitting of a prior distribution employing generic data. Some methods for reducing or combining such data in fitting a prior are discussed. Subsequently, several classes of priors that have been found useful in applications are introduced. Particular note is given to the class of noninformative prior distributions, useful when there are few or no satisfactory prior generic, including subjective, data. Log-normal, gamma, and beta prior distributions are presented for possible use when usable prior generic data are available. A form of a gamma also applies in the noninformative case.

6.3.3.1 Sources of Data for Use in Bayesian Estimation

As was noted in Section 6.1, three types of information about the reliability parameter of interest are often available: (1) engineering knowledge about the design, construction, and performance of the component; (2) the past performance of similar components in similar environments; and (3) the past performance of the specific component in question. The first two types constitute the "generic" information (or data) and may include varying degrees of subjective judgment. The third type, constituted of objective data, is the "application-specific" information (or data). Generic data are employed in the development of a prior distribution. The application-specific data are used to revise the prior into a posterior distribution. Commonly applied prior distributions are next described.

6.3.3.2 Noninformative Prior Distributions

"Noninformative" prior distributions are a class of priors that minimize the relative importance of the prior compared with that of the experience data in generating a posterior estimate. There are many ways of precisely quantifying this basic notion and hence a variety of classes of noninformative priors and corresponding methods for their attainment in practice. The concept employed here for the noninformative prior is that of Martz and Waller (1982) (adopted from Box and Tiao, 1973) in which, roughly speaking, a prior is said to be noninformative if the application-specific data serve only to change the location of the corresponding likelihood function and not its shape. This and other ideas have also been discussed by Jeffreys (1961), and a summary of the relevant literature on this subject has been presented by Parry and Winter (1981).

Noninformative priors are useful when little or no generic prior information is available; they should not be used when there is such information, because they deliberately downgrade its role in the estimation process. Frequently, Bayesian posterior estimates from noninformative priors are close to the corresponding classical estimates when the latter are also available, a fact illustrating the versatility of the Bayesian method. The noninformative priors for failure-on-demand probabilities and for failure rates, respectively, are discussed below.

6.3.3.3 Natural Conjugate Prior Distributions

Natural conjugate prior distributions have the property that, for a given likelihood function, the posterior and prior distributions are members of the same family of distributions. In such cases, the posterior distribution has a closed-form analytical representation when the prior does, which is of the same character as the prior's except for the values of its parameters, and accordingly the expressions for computing the Bayesian point and interval estimates can usually be represented in terms of well-defined probabilities. The parameters of such priors and posteriors are often especially easy to interpret, playing the role of failure data entirely analogous to the particular data used in the likelihood function. This will be illustrated below. Such families of priors are often rich enough and flexible enough to permit the analyst to model reasonably a wide range of prior data that may be encountered (Martz and Waller, 1982). Finally, there are well-developed methods for fitting natural conjugate priors to generic prior data.

For these reasons, the use of natural conjugate priors is desirable* whenever the exact form of the prior has not already been specified as part of the generic prior information, but the prior data are sufficient to determine a reasonable member of the natural conjugate family. If incomplete information exists on the prior, as often happens, the analyst will have the flexibility to select the form of the distribution, and the conjugate prior is often the natural selection. However, a sensitivity analysis should be performed to confirm this choice.

6.3.3.4 Using Generic Data Sources

The generic prior data must be reduced to a form that permits the selection of a specific prior distribution from a suitable family. For example, if a log-normal family has been selected, the two log-normal parameters must be determined from the generic data. If there are multiple sets of generic prior data, these must likewise be reduced to a common consensus prior by one of several means.

A Single Source

For convenience consider the case of failure rate (per unit time) estimation. If a two-parameter prior distribution is to be fitted, such as a log-normal or a gamma distribution, the generic data must contain at least two independent pieces of information. For example, the generic data may consist of upper and lower percentile limits on the failure rate. Each of these limits is then equated to its theoretical counterpart derived from the prior family considered. Since each theoretical expression will be a function of the two prior parameters, the two equations can be solved simultaneously for the values of the two parameters.

* Nevertheless, a log-normal prior is often used, despite the fact that it is not naturally conjugate in relation to binomial or Poisson experience data, because of the convenient multiplicative property of the log-normal distribution: a product of log-normals is log-normal. This facilitates the calculation of mincutset probabilities as products of basic event probabilities (recall Chapter 4).

Example 1. Given that a diesel generator starts successfully, its subsequent hourly failure rate is given in the Reactor Safety Study (Nuclear Regulatory Commission, 1975) as a log-normal distribution with 5th percentile $\lambda_L = 3 \times 10^{-4}$ and 95th percentile $\lambda_U = 3 \times 10^{-2}$. For the log-normal distribution the following pair of equations applies:

$$\Phi \left[\frac{\ln (3 \times 10^{-4}) - \xi}{\sigma} \right] = 0.05$$

and (6-15)

$$\Phi \left[\frac{\ln (3 \times 10^{-2}) - \xi}{\sigma} \right] = 0.95$$

where ξ and σ are parameters of the log-normal family and Φ (.) is the standard normal cumulative distribution function. Since $\Phi(-1.645) = .05$ and $\Phi(1.645) = 0.95$,

$$\ln (3 \times 10^{-4}) - \xi = -1.645\ \sigma$$

and (6-16)

$$\ln (3 \times 10^{-2}) - \xi = 1.645\ \sigma$$

from which $\xi = -5.81$ and $\sigma = 1.40$. Thus, the fitted log-normal prior becomes

$$f(\lambda) = \frac{1}{1.40\ \lambda \sqrt{2\pi}} \exp \left[- \frac{1}{2\ (1.40)^2} (\ln \lambda + 5.81)^2 \right] \quad (0 < \lambda < \infty) \qquad (6\text{-}17)$$

An alternative technique is considered later.

Similar techniques can be used for generic data such as means or medians. However, if only a "best" point estimate is given, there will usually be a need for some additional specification by the analyst (or the two-parameter log-normal might be replaced by a one-parameter distribution). First, he must decide whether to use the mean, median, or mode of the distribution as the suitable central value representing the "best" estimate. Second, the analyst may have to introduce a second parameter value in order to define a distribution without ambiguity. For example, suppose one is to fit a gamma prior for a failure rate when the only available datum is the mean of the generic rate. Since the mean does not uniquely determine a gamma distribution, the variance could also be introduced and treated as an unspecified parameter (or as noted in Section 6.3.5.3 below, the exponential, which is a one-parameter gamma with the other parameter set equal to 1, might be used as the best representation of the limited prior information actually available).

Sometimes the prior data from a single generic source are inconsistent in the sense that no common prior distribution can be fitted to the data. There is no universally accepted method of rectifying such inconsistencies, but any of several approaches could be taken. One would be to take the set of all priors implied by the generic data and define some "most conservative" criterion to select a single prior from the set. Another would be to consider the entire set of priors as representing multiple sources of generic data and employ the procedures suggested in the discussion that follows.

6.3.3.5 Combining Multiple Sources

Often, generic prior data from multiple sources must be reduced to a single prior distribution that satisfactorily reflects and incorporates the views of each source. The multiple sources might be generic data from two or more studies that report on the same generic component; they may consist of the opinions of several experts about the same component; or, as noted above, the multiple "sources" may consist of the set of unrectified priors obtained from a single inconsistent source.

A number of procedures are suggested here for forming a consensus prior distribution. For convenience, consider a failure rate estimation as before. If each source provides both a point and an interval estimate, the first method is to pool (combine) the estimates by means of simple geometric averaging techniques:

$$\hat{\lambda} = \left(\prod_{i=1}^{n} \hat{\lambda}_i\right)^{1/n} \tag{6-18}$$

This is equivalent in effect to forming the usual arithmetic average of failure rates described by their logarithms. This estimate implicitly assumes that the underlying sources are statistically independent and of equal importance. If the sources are unequal in their contribution to the consensus prior, a weighted geometric mean could be used with the uniform 1/n factor replaced by weights chosen to reflect the importance of, or confidence in, each source (see, e.g., Fragola et al., 1987).

Martz and Bryson (1982) have developed a classical statistical model for combining multiple sources of data. The resultant maximum likelihood consensus point estimator is a weighted geometric mean of the individual estimates in which the weights are simple functions of the uncertainty bounds supplied by each data source. A corresponding consensus confidence interval estimator is also provided. The maximum likelihood point estimator reduces to Equation 6-18 under two conditions: if each data source reports the exact same range of uncertainty, and if there is no location bias in the individual estimates.

The foregoing pooling methods have been shown to provide good point estimates; however, the combined interval estimates generally have tended to be too narrow and thus have had less than the desired assurance.

The second method yields a consensus prior that is generally more diffuse (spread out) than that obtained by the method just described. This method, discussed, e.g., by Winkler (1968), is often referred to as the "mixture method." It involves fitting a suitable prior to each generic source and then combining the individual prior distributions by forming a mixture,

$$f(\lambda) = \sum_{i=1}^{n} w_i f_i(\lambda) \tag{6-19}$$

The coefficients w_i are positive weights that sum to 1. Winkler (1968) suggests several methods for determining the weights. In the absence of any reason for preferring one source over another, the selection $w_i = n^{-1}$ is an obvious possibility. An interesting feature of this method is that it may yield a non-unimodal prior distribution. If such a mixture is used as a prior distribution, the corresponding posterior distribution from Equation 6-9 will also be a mixture of the individual (component) posterior distributions, namely,

$$f(\lambda \mid E) = \sum_{i=1}^{n} w_i' f_i(\lambda \mid E) \tag{6-20}$$

where the new (updated) weights are

$$w_i{}' = \frac{w_i \int_0^{\infty} f_i\,(\lambda)\ L\ (E|\lambda\ )\,d\lambda}{\sum_{i=1}^{n} w_i \int_0^{\infty} f_i\,(\lambda)\ L\ (E|\lambda\ )\,d\lambda} \quad (i = 1, \ldots, n) \tag{6-21}$$

Since this method generally yields a more diffuse consensus prior than does geometric averaging, it provides more conservative interval estimates. For this reason it is often preferred. However, it should be pointed out that the mixture method is computationally more difficult; numerical techniques are frequently required for determining such quantities as the prior moments and percentiles.

A third method, called a "two-stage" Bayesian procedure by Kaplan (1981 a), uses a Bayesian procedure for forming the prior (stage 1) before combining the prior with the likelihood function (stage 2). To describe the two-stage method, assume that the problem to be solved is to estimate the failure rate of machine S and express the degree of confidence in this failure rate, given the following relevant information:

E_1: engineering knowledge of the design and construction of the machine
E_2: past performance of similar machines in similar applications
E_3: past performance of the specific machine in question

The information E_3 is of the form

$$E_3 = < h_s, T_s > \tag{6-22}$$

that is, a doublet stating that machine S has failed h_s times in T_s years. This information is used in Bayes' theorem:

$$f\,(\lambda \mid E_1,E_2,E_3\,) = \frac{f\,(\lambda \mid E_1,E_2)\ L(\,E_3 \mid \lambda,\,E_1,E_2)}{\int_0^{\infty}\ f\,(\lambda \mid E_1,E_2)\ L(\,E_3 \mid \lambda,\,E_1,E_2)\ d\lambda} \tag{6-23}$$

where $f(\lambda \mid E_1,E_2,E_3)$ is the posterior probability distribution for λ. This distribution expresses the final state of knowledge about λ in light of all the evidence E_1,E_2, and E_3. On the right, $f(\lambda \mid E_1,E_2)$ is the "prior" distribution representing the state of knowledge without information E_3 but including E_1 and E_2.

This use of Bayes' theorem to incorporate the specific evidence E_3 is a conventional application of Bayes' theorem and is the second stage of the two-stage approach. The first stage of the two-stage approach is aimed at determining the prior $f(\lambda \mid E_1,E_2)$, from the information E_2, which is of the form

$$E_2 = \{ <h_1,T_1>, \ldots , <h_M,T_M> \} \qquad \textbf{(6-24)}$$

E_2 then is the set of doublets giving the operating experience of a set of M components deemed similar to the component being analyzed.

To use E_2, this set of M components is thought of as a sample from an infinite population Q of similar components. Considering the whole of Q, there is a frequency distribution $\Phi(\lambda)$, where λ is the failure rate of a member of Q, such that $\Phi(\lambda)\,d\lambda$ is the fraction of the population with failure rates in the interval $d\lambda$. $\Phi(\lambda)$ is called the "population variability curve" for the population Q.

If the population variability curve were in fact known, it could be used as a prior, that is,

$$f(\Phi \mid E_1,E_2) = \Phi(\lambda) \qquad \textbf{(6-25)}$$

Since $\Phi(\lambda)$ is not known, it is necessary to express what is known or can be inferred about $\Phi(\lambda)$ from the evidence E_2. For this purpose, consider the function $\Phi(\lambda)$ as being embedded in a space of functions $\{ \Phi(\lambda) \}$. Then a probability distribution, call it $f(\Phi \mid E_1,E_2)$ over this space F of functions exists, expressing knowledge of where, in F, $\Phi(\lambda)$ is located. For this purpose, write the "first-stage" application of Bayes' theorem in the form

$$f(\Phi \mid E_1,E_2) = \frac{f(\Phi \mid E_1)\, L(E_2 \mid \Phi, E_1)}{\int_0^\infty f(\Phi \mid E_1)\, L(E_2 \mid \Phi, E_1)\, d\Phi} \qquad \textbf{(6-26)}$$

Thus $f(\Phi \mid E_1, E_2)$ is the state of knowledge about Φ "posterior" to having the information E_2.

Once $f(\Phi \mid E_1, E_2)$ is known, the desired prior $f(\lambda \mid E_1, E_2)$ for the second stage of the process can be calculated from

$$f(\lambda \mid E_1, E_2) = \int_F f(\Phi \mid E_1, E_2)\, d\Phi \qquad (6\text{-}27)$$

Kaplan (1981b) uses "discretization" techniques to find the population-variability curve. This can be illustrated by choosing a two-parameter family of log-normal* curves as follows:

$$\Phi_{ij}(\lambda) = \frac{1}{\sqrt{2\pi}\,\lambda\sigma_j} \exp\left\{ - \frac{[\ln(\lambda/\mu_i)]^2}{2\sigma_j^2} \right\} \qquad (6\text{-}28)$$

where the two parameters μ_i, σ_j range over a discrete "grid." Thus,

$$p(\Phi_{ij} \mid E_1, E_2) = \frac{p(\Phi_{ij} \mid E_1)\, p(E_2 \mid \Phi_{ij}, E_1)}{\sum_{i=1}^{I} \sum_{j=1}^{J} p(\Phi_{ij} \mid E_1)\, p(E_2 \mid \Phi_{ij}, E_1)} \qquad (6\text{-}29)$$

and

$$p(E_2 \mid \Phi_{ij}, E_1) = \sum_{m=1}^{M} \left[\int_0^\infty \Phi_{ij}(\lambda)\, \frac{(\lambda T_m)^{K_m}}{K_m!} \exp(-\lambda T_m)\, d\lambda \right] \qquad (6\text{-}30)$$

where M is the number of components m each with data K_m failures in T_m hours.

* The use of this family of log-normal curves is illustrative. Any desired family of curves could be used, subject only to the requirement that somewhere in the family there would be at least one good approximation of the true variability curve Φ.

The prior p ($\Phi_{ij} \mid E_i$) is the information that describes the grid of the parameters μ_i and σ_j. This is determined from experience, or it could be a noninformative prior.

A further simplification can be made by finding a "best estimate" for Φ_{ij} , or the mean value for the distribution p ($\Phi_{ij} \mid E_1,E_2$); that is,

$$\Phi(\lambda) = \sum_{i,j} \Phi_{ij}(\lambda)\, p(\Phi_{ij} \mid E_1, E_2) \qquad \textbf{(6-31)}$$

This could then become the final prior for combining with the likelihood function from E_3.

6.3.3.6 Using Expert Opinion

The opinions of Subject Matter Experts (SMEs) are often used for a prior probability distribution when other information is inadequate. If neither physical nor theoretical models are available and relative frequency is unavailable as well, subjective assessment is the only alternative for obtaining a probability. The practical feasibility of this alternative is supported not only by theoretical foundations that show properly structured judgments about uncertain events can be expressed as probabilities through practical assessment procedures. Holloway (1979) reviews the basis for these procedures and gives examples for several assessment approaches. The following summary of assessment procedures draws on his book. After this summary, well-known cautions and guidelines for interpreting and reviewing expert opinions are presented to highlight the care and caveats that must accompany judgmental quantitative assessment.

However, the user of this handbook should be cautioned against the indiscreet use of the methods described in this section. These techniques and results are not all necessarily applicable to probabilistic risk assessments, that often treat extremely small probabilities of various events. More research is needed to determine the full applicability of these methods and findings to PRAs. The user should be aware that the subjective estimates frequently used in PRAs can have large biases and errors (see, e.g., Fragola et al., 1987; Tversky and Kahneman, 1974; Kahneman and Tversky, 1979).

Assessment Lotteries

An assessment lottery is a physical example of a random process. The uncertainty represented by the lottery must be easily recognized by the expert and have a definite, objective probability. Such a lottery is the reference scale that measures an expert's degree of belief about the uncertain event. The operational definition for subjective probability, then, is the fraction of this reference uncertainty scale that makes an expert just indifferent between the assessment lottery and the feeling of uncertainty toward the event being assessed.

One example of an assessment lottery is implemented with an urn containing balls of different colors, some fraction being one color and the rest the other color. Drawing a ball at random from the urn is supposed to provide a visualization of an objective probability. Spetzler and Stael von Holstein (1975) developed and clinically tested a similar procedure that uses the spinning of a reference wheel as the assessment lottery device. Their experience has shown that these probability wheels provide a strong visual image of an uncertain process.

Assessment Procedures

Two general approaches to subjective probability assessment are in practical use, either the direct approach or the indirect approach. With the direct approach, the expert is asked to declare the probability number associated with his feeling of uncertainty for the occurrence of an event. With the indirect approach, an expert is asked to choose between a reference assessment lottery and his uncertain feeling (his degree of belief) in an opinion or judgment. Until an expert has shown an ability both to form a knowledgeable opinion and to assess, unaided, a probability for his degree of belief associated with that opinion, the indirect approach is preferred. The well-known difficulties in obtaining useful subjective probability assessments are summarized below in the section entitled "Validity of Expert Opinion." These difficulties are magnified in inexperienced, unaided direct assessments. The references in that section describe some experience in comparing the two approaches.

The direct approach has the expert state a number that represents the assessment of the probability. Some studies have shown it possible for people to become better at assessing their own feelings of uncertainty as probabilities (see, for example, Stael von Holstein, 1970). This improvement in direct assessment comes from specific training and guided practiced discipline rather than by trial and error. A good direct assessment comes from one who is both an experienced expert in what is known about a technical area (as well as how much is not known) and an experienced expert on how to express that judgment with little cognitive bias. This combination of kinds of expertise is uncommon .

Assessment lotteries are used in the indirect approach to disclose the subjective probability. This external reference is used as a scale to measure the internal degree of belief an expert holds for an opinion. Dividing between the expert and the assessors the responsibility to provide both a well-founded, knowledgeable judgment and an accurate representation of that judgment as a probability allows the use of expert opinion in PRAs. Most technical experts are not practiced, good probability assessors in themselves. Using the indirect approach improves the quality of expert opinion over that obtained by unaided, inexperienced direct assessment. However, Fischhoff et al. (1981) have shown that people qualified as technical experts are by no means necessarily qualified as probability assessors of that expertise. Lindley and Singpurwalla (1986) describe a mathematically complete procedure for introducing the assessor's evaluations of the experts' opinions into the final probability assessments.

A second indirect subjective probability approach may be applicable when a set of events with comparable probabilities ranging from reasonably high to possibly very low can be considered. The idea is to elicit from the experts the ratios of successive pairs of probabilities, starting with the largest one and in order of decreasing values. Thus, if the succession of probabilities is P_1, P_2, P_3,..., etc., what is elicited is how much smaller P_2 is than P_1, P_3 is than P_2, etc. Such relative assessments may be much easier than absolute ones, and the smallest probabilities, which generally would be all but impossible to estimate absolutely, now may be implied fairly simply as reasonable fractions of the somewhat larger probabilities just above in the sequence. Finally, to transform all the relative probabilities into absolute ones requires one absolute probability assessment. This can be the estimated value of the largest probability P_1, whose magnitude should make a subjective assessment relatively easy. Of course, early estimation errors will propagate down the sequence and so care must be taken with this procedure to verify its results wherever possible. One or two additional absolute assessments of members of the sequence would be of great help in this. The sequence should be selected to facilitate this, if possible with elements included which have already available or relatively easily assessed independent probability estimates.

Assessment Models

The representation used to model the uncertain event, either intuitively or formally, is a significant part of obtaining a good assessment. How the SME thinks about the problem of giving a judgment on the event likelihood should be recorded (see the discussion of "Recording Expert Opinion," below). It is this representation that fashions the eventual probability that is assessed. If disputes or questions arise in reviewing the quality of the expert opinion, a brief description of the thought model can focus the issue on a particular facet of that judgment.

Often, the expert is better able to provide a judgment by refining the event description into underlying events or factors. This formal assessment model can be subdivided until the expert finds it easy to examine each part, provide an opinion conditioned on each one, and review the formally computed probability of the original event for completeness and accuracy. This aid to assessment relieves an expert from making logical, or procedural, errors in combining the underlying knowledge. Reducing this source of error with the use of assessment models allows the assessor to focus on revealing a more subtle bias in the judgment.

Validity of Expert Opinion

The validity of a subjective assessment comes from two distinct parts: the knowledge content provided by the SME and the procedural process provided by the assessor. If the expert is playing both of these roles, the distinction blurs, but it is still useful to describe the source of inaccuracies.

The content factor is evaluated from the credentials provided by the expert. Identifying who knows what and how much is a routine task for a professional community. Even for a recognized expert, a peer review can use the assessment model to judge whether or not all the significant factors were included in the expert's opinion. Inaccuracies, disputes, omissions, and limits to knowledge can then be examined to improve the accuracy of the substantive, or content, portion of the probability assessment.

The procedural process is more difficult to evaluate. The judgmental processes used by the expert, the effect the assessor has on expanding or limiting the formation of the expert's opinion, the effect of misunderstandings, and the natural cognitive limits on human information processes are all hidden factors in a practical assessment. Clinical studies, however, have examined these process factors that affect expert opinion. These studies provide a catalog of possible sources of inaccuracy due to bias and the extent of their effects.

It is well known that various biases may accompany the subjectively quantified assessments of an expert. For example, Alpert and Raiffa (unpublished work, 1969) found that experts often overestimate the degree of certainty of their estimates and claim too high a level of assurance. They observed that interval estimates for which 98-percent assurance was claimed tended in reality to have an assurance of about 70 percent (i.e., to include the correct value 70 percent of the time). Alternatively stated, interval estimates are often too narrow for the assurance level that is claimed. Tversky and Kahneman (1974) attribute such bias in part to a phenomenon of "anchoring": the expert tends to focus, or "anchor," on an initial guess and is reluctant to deviate too far from that guess in accounting for possible misjudgment. The results of such studies suggest that the assurance associated with expert-supplied interval estimates should be reduced from that claimed. For example, if a 90-percent interval estimate is solicited, then the interval could perhaps be considered to be an actual 70-percent interval in fitting a prior.

It is also well known that the manner chosen to encode (solicit) the subjective probabilities held by the expert is crucial and may significantly affect the quality of the information. Spetzler and Stael von Holstein (1975) describe and recommend a structured-interview procedure and suggest a number of techniques for reducing biases in the quantification of judgment.

Holloway (1979) finds two results of these studies encouraging. First, persons who are procedural experts in obtaining probability distributions are able, by using a variety of assessment techniques, to elicit consistent, well-founded judgments from substantive experts. Second, the substantive experts who are knowledgeable about the event being assessed are able to learn quickly about the significant procedural factors of probability assessment.

Recording Expert Opinion

The procedure used for assessing expert opinion and the assessment model used by the expert to construct the judgment should be described in a record of the expert opinion.

A subjective probability estimation is an evaluation. The important procedural and substantive factors in that evaluation should be recorded, like any other engineering analysis, to permit a peer review to determine the quality of that result.

This record does not have a standard format; however, with time and experience, one may evolve. Nevertheless, the probability number can be meaningless without a description of how it was obtained and what were its principal foundations.

6.3.3.7 Beta Prior Distributions

The beta family of prior distributions is the conjugate family when failure-on-demand probabilities are estimated with a binomial likelihood function (see Section 6.3.4.1, below). To fit a beta prior, values of the two prior beta parameters must be selected.

Martz and Waller (1982) present a table-lookup procedure, along with two sets of tables, that can be directly used to determine the beta parameter values. Two situations are considered: (1) when the prior mean and 5th percentile of the prior distribution of failure-on-demand probabilities are specified and (2) when the prior mean and 95th percentile are specified. The procedure then yields directly the two beta parameters, as described by Martz and Waller with examples.

Mosleh and Apostolakis (1982) also describe a procedure for determining the beta parameter values corresponding to various combinations of 5th, 50th, and 95th percentiles as well as the mean. Their procedure is to approximate the beta distribution as a gamma distribution and use corresponding techniques for determining the gamma parameters.

Ahmed *et al.* (1981) have developed a computer code, called BURD, that finds the beta parameter values corresponding to specified 5th and 95th percentile values.

6.3.3.8 Gamma Prior Distributions

The gamma family of prior distributions is the conjugate family when failure rates are estimated with a Poisson likelihood function (see Section 6.2.5.3). The gamma family is a two-parameter (α, β) family, and both shape (α) and scale (β) parameter values must be identified by specifying some two conditions.

Martz and Waller (1982) present a simple procedure for determining the values of both parameters when two percentiles are given, corresponding to tail areas of 0.5, 1, 2.5, 5, 10, 25, 50, 75, 90, 95, 97.5, 99, or 99.5 percent. Mosleh and Apostolakis (1982) also present a procedure for determining the two gamma parameter values for specified pairs of values--the (5th, 95th), (5th, 50th), (50th, 95th), (mean, 5th), or (mean, 95th). (However, caution must be exercised on the use of the mean with a percentile, as will be noted below.) Ahmed et al. (1981) describe the use of the BURD code to determine the gamma parameter values for specified 5th and 95th percentile values.

It may sometimes be the case that prior information is available only on one parameter, some percentile or the mean. The conjugate properties of the gamma, as well as its mathematical convenience in some component-level to system-level computations, can be retained if the special gamma, the exponential, which is gamma with $\alpha = 1$, is employed as the prior.* If no prior information at all is available, and the noninformative prior discussed above is used, it is to be noted that Poisson sample data then lead again to a gamma posterior. Thus, in all these cases, as well as the general one in which two prior percentiles are available, a gamma posterior results. For computational purposes it is worth recognizing that a gamma whose α is a multiple of 1/2 is equal to a weighted X^2, and this in turn can be numerically evaluated easily with the Wilson-Hilferty normal approximation (Mann, Schafer, and Singpurwalla, 1974). It should be possible in most practical applications to satisfactorily approximate the prior value of α by the nearest multiple of 1/2 (or to be conservative, the next higher multiple of 1/2), and thus enable the use of a X^2 as the prior. The posterior of course will then be another X^2 if Poisson sample data are applied. In Section 5.4, this will be shown to facilitate the development of certain system-level uncertainty distributions.

As indicated earlier, a note needs to be made about the use of estimates of one percentile and the mean in establishing a gamma prior. It will be recognized first that in general these values cannot uniquely define a distribution of a given family because they do not specify how much probability will exist between the mean and the percentile -- in principle, the

* As with all assumptions of a prior, a "pre-posterior" simulation can be carried out of the implications of this choice when various possible future data samples are combined with the prior to generate possible posteriors. If the results seem inappropriate to the expert providing the single mean or percentile that produced the exponential, he may be inclined to consider further a second parameter value that would lead to more satisfactory posterior results.

mean could be close in probability to any percentile, depending on the skewness. Second, the two given values may be inconsistent with any possible range of values of the probability in any gamma (or other particular) distribution between the mean and the given percentile. Thus, there may be no gamma distribution (other than some good or bad approximation) consistent with the given mean and percentile estimates. For further details on this point and other aspects of the development, approximation, and use of the gamma distribution for the uncertainty in component failure rates, see Heubach and Philipson (1985, 1986).

As already noted, when employing any of the foregoing priors in a particular case it will often be valuable to conduct a "pre-posterior" sensitivity analysis. This tests whether the experts' opinions or other generic data leading to a prior are consistent with reasonable expectations of possible future experience. Hypothetical Poisson samples can be input to revising the prior to produce posterior means and percentiles. If the values arrived at seem out of line given the hypothesized data, then the prior can be reassessed. The prior's suitability can also be tested by evaluating the probability of observing each of some range of reasonably expectable future samples, given the prior distribution. If the probability values arrived at are too low, the distribution should be rejected. Note that, in principle, a prior could be "designed" in this way, deriving one that is most suited to reasonable futures, instead of deriving it in accordance with preexisting data.

6.3.3.9 Log-normal Prior Distributions

The log-normal distribution is frequently used as a prior distribution for failure rates, especially when the failure rates typically encountered are so low (say, 10^{-6} per demand or per unit time) as to make a logarithmic transformation attractive. A simple procedure follows for determining the log-normal parameters ξ and σ (see also Sections 6.3.4.4, 6.3.5.4).

Suppose that two symmetrically located percentiles are specified for the log-normal, denoted by λ_γ and $\lambda_{1-\gamma}$, where $0 < \gamma < 0.5$. Thus,

$$P(\lambda < \lambda_\gamma) = P(\lambda > \lambda_{1-\gamma}) = \gamma \qquad \textbf{(6-32)}$$

The geometric mean of the percentiles is defined as

$$M = (\lambda_\gamma \lambda_{1-\gamma})^{1/2} \tag{6-33}$$

and a generalized error factor is

$$EF = (\lambda_{1-\gamma} / \lambda_\gamma)^{1/2} \tag{6-34}$$

Then the desired parameter values are

$$\xi = \ln M \, , \sigma = \ln EF / Z_{1-\gamma} \tag{6-35}$$

where $Z_{1-\gamma}$ is the 100(1 - γ)th percentile of a standard normal distribution. In this case the mean, the variance, the mode, and the median of the fitted log-normal distribution can be found from the parameters as follows:

$$\begin{aligned} &\text{Mean:} && \exp (\xi + \sigma^2/2) \\ &\text{Mode:} && \exp (\xi - \sigma^2) \\ &\text{Median:} && \exp (\xi) \equiv M \\ &\text{Variance:} && [\exp (2 \xi + \sigma^2)] [\exp (\sigma^2) -1] \end{aligned} \tag{6-36}$$

It is further observed that M is the median of the log-normal distribution and that the two percentiles are $\lambda_{1-\gamma}$ = (EF)(M) and λ_γ = M/(EF), in accord with the notion of an error factor.

Example. Let $\lambda_{.05} = 3 \times 10^{-4}$ and $\lambda_{.95} = 3 \times 10^{-2}$. Then $M = 3 \times 10^{-3}$ and EF = 10. These are then substituted into Equations 6-35 to obtain ξ = -5.81 and σ = 1.40, for the latter making use of the fact that $Z_{.95}$ = 1.645. Equations 6-36 give for the mean, mode, median, and variance the values 8×10^{-3}, 4×10^{-4}, 3×10^{-3}, and 4×10^{-4}, respectively.

6.3.4 Estimating Failure-on-Demand Probabilities

6.3.4.1 Binomial Likelihood Function

As has been noted, the binomial distribution is the distribution of the number of failures, r, in n independent demands, in each of which the component has a constant failure-on-demand probability p. Given this statistical framework, the likelihood in Equation 6-9 is the binomial distribution, given by

$$L(E|p) = \frac{n!}{r!(n-r)!} p^r (1-p)^{n-r} \tag{6-37}$$

for r = 0,1,2,...,n and the parameter p between 0 and 1. If the parameter p is small and n is sufficiently large (e.g., p < .1 and np > 5), then Equation 6-37 will usually be most conveniently approximated by the Poisson distribution:

$$L(E|p) = (np)^r \exp(-np) / r! \tag{6-38}$$

Because in all cases of interest the number of demands is large in comparison with the number of failures, r can be treated as being able to assume any nonnegative integral value with the larger values of r contributing negligibly to the probability distribution.

In the Bayesian approach, the parameter p is regarded as a random variable due to uncertainty, with a specified prior distribution. Three methods of generating a prior for p are considered: (1) a noninformative prior; (2) a natural conjugate beta prior; and (3) a log-normal prior. Only the major results and formulas required to compute appropriate moments and estimates are given here. Details can be found in Martz and Waller (1982).

6.3.4.2 Noninformative Prior Distribution

A noninformative prior density is

$$\left[p(1-p)\right]^{-0.5} / \pi \quad (0 < p < 1) \tag{6-39}$$

The prior mean, median, and variance are as follows:

Prior mean:	0.5	
Prior median:	0.5	**(6-40)**
Prior variance:	0.125	

and the prior 100 (1 - γ)% symmetric probability interval is

(6-41)

$$\left(\frac{0.5}{0.5 + 0.5\ F_{1-\gamma/2}(1,1)} , \frac{0.5\ F_{1-\gamma/2}(1,1)}{0.5 + 0.5\ F_{1-\gamma/2}(1,1)} \right)$$

where $F_{1-\gamma}$ (a,b) is the 100 (1 - γ)th percentile of an F-distribution with a and b degrees of freedom.

The posterior density, after observing r failures in n demands, is

(6-42)

$$\frac{\Gamma(n+1)}{\Gamma(r+0.5)\ \Gamma(n-r+0.5)}\ p^{r-0.5}(1-p)^{n-r-0.5} \quad (0 \leq p \leq 1)$$

and the formulas for calculating the posterior mean, median, and density are:

Posterior mean:

$$(r+0.5)/(n+1)$$

Posterior median: **(6-43)**

$$\frac{r+0.5}{r+0.5+(n-r+0.5)\,F_{0.5}(2n-2r+1, 2r+1)}$$

Posterior variance:

$$\frac{(r+0.5)(n-r+0.5)}{(n+1)^2\ (n+2)}$$

The posterior 100(1 - γ)% symmetric probability interval is

(6-44)

$$\left(\frac{r + 0.5}{r + 0.5 + (n - r + 0.5)\, F_{1-\gamma/2}\,(2n - 2r + 1, 2r + 1)}, \right.$$
$$\left. \frac{(r + 0.5)\, F_{1-\gamma/2}\,(2r + 1, 2n - 2r + 1)}{n - r + 0.5 + (r + 0.5)\, F_{1-\gamma/2}\,(2r + 1, 2n - 2r + 1)} \right)$$

6.3.4.3 Beta Prior Distribution

For the beta prior distribution, the prior density is

(6-45)

$$\frac{\Gamma(n_0)}{\Gamma(r_0)\Gamma(n_0 - r_0)}\, p^{r_0 - 1}\,(1 - p)^{n_0 - r_0 - 1} \qquad (0 \le p \le 1)$$

where the positive values n_0 and r_0, parameters of the beta distribution, may be interpreted as the numbers of equivalent demands and failures, respectively, in the prior data. The prior mean, median, and variance are:

Prior mean: r_0 / n_0

Prior median: (6-46)

$$\frac{r_0}{r_0 + (n_0 - r_0)\, F_{0.5}\,(2\, n_0 - r_0, 2 r_0)}$$

Prior variance:

$$\frac{r_0\,(n_0 - r_0)}{n_0^2\,(n_0 + 1)}$$

and the prior 100(1 - γ)% symmetric probability interval is

(6-47)

$$\left(\frac{r_0}{r_0 + (n_0 - r_0) F_{1-\gamma/2} (2n_0 - 2r_0, 2r_0)} , \quad \frac{r_0 \; F_{1-\gamma/2} \; (2r_0 , 2n_0 - 2 r_0)}{n_0 \; - (n_0 + r_0) F_{1-\gamma/2} \; (2r_0 , 2n_0 - 2 r_0)} \right)$$

The posterior density is given by

(6-48)

$$\frac{\Gamma(n + n_0)}{\Gamma(r + r_0) \; \Gamma(n - r + n_0 - r_0)} \; p^{r + r_0 - 1} \; (1 - p)^{n - r + n_0 - r_0 - 1} \qquad (0 \le p \le 1)$$

and the other formulas are as follows:

Posterior mean:

$$(r + r_0) / (n + n_0)$$

Posterior median: **(6-49)**

$$\frac{r + r_0}{r + r_0 + (n - r + n_0 - r_0) F_{0.5} \; (2n - 2r + 2\, n_0 - 2\, r_0 , 2\, n_0 + 2\, r_0)}$$

Posterior variance:

$$\frac{(r + r_0) \; (n - r + n_0 - r_0)}{(n + n_0)^2 \; (n + n_0 + 1)}$$

Posterior 100(1 - γ)% symmetric probability interval:

(6-50)

$$\left(\frac{r+r_0}{r+r_0+(n-r+n_0-r_0)F_{1-\gamma/2}(2n-2r+2n_0-2r_0\,,2r+2r_0)}\right.$$

$$\left.\frac{(r+r_0)F_{1-\gamma/2}(2r+2r_0,\ 2n-2r+2n_0-2r_0)}{n-r+n_0-r_0+(r+r_0)F_{1-\gamma/2}(2r+2r_0,\ 2n-2r+2n_0-2r_0)}\right)$$

6.3.4.4 Log-normal Prior Distribution

As noted earlier, the log-normal distribution is often used as a prior distribution for the uncertainty in p, but its parameters must be so chosen that the probability density outside the actual range of p--that is, above the value $p = 1$ — is sufficiently small to be ignored or effectively truncated.

The prior density is

(6-51)

$$\frac{1}{\sigma\, p\sqrt{2\pi}}\ \exp\left[-\frac{1}{2\sigma}\ (\ln p-\xi)^2\right]\quad (p > o)$$

The prior 100(1 - γ)% symmetric probability interval is:

(6-52)

$$\left(\exp(\xi - Z_{1-\gamma/2}\sigma),\ \exp(\xi + Z_{1-\gamma/2}\ \sigma)\right)$$

The posterior distribution cannot be obtained in closed form. However, the approximation given in Equation 6-10 can be used to approximate the posterior distribution where $f(p_i)$ denotes the area under the log-normal prior over an interval represented by $p = p_i$ and $L(E \mid p_i)$ denotes either Equation 6-37 or 6-38 evaluated at $p = p_i$ for the selected set of discrete values p_i $(i = 1,2,\ldots,m)$.

6.3.5 Estimating Constant Failure Rates

6.3.5.1 Poisson Likelihood Function

As discussed above, a common assumption in reliability models is that failure times are independent, with a common exponential (constant failure rate) distribution. It follows that the distribution of the number of failures r in a fixed total operating time T has a Poisson distribution. In this case the likelihood function defined in Equation 6-9 is the Poisson density given by:

$$L (E \mid \lambda) = (\lambda T)^{r} \exp (-\lambda T) / r! \quad (r = 0, 1,2,...) \tag{6-53}$$

where λ denotes the constant failure rate.

Three cases again are considered: (1) a noninformative prior distribution, (2) a natural conjugate gamma prior distribution, and (3) a log-normal prior distribution for λ (see Martz and Waller, 1982).

6.3.5.2 Noninformative Prior Distribution

The various formulas for the noninformative prior distribution are:

Prior density: $\sim\lambda^{-1/2}$ (an improper distribution) $(\lambda > 0)$ **(6-54)**

Posterior density: $$\frac{T^{r+1/2}}{\Gamma (r + 1/2)} \lambda^{r-1/2} \exp (- \lambda T) \tag{6-55}$$

Posterior mean: $(2r + 1) / (2T)$

Posterior median: $X^2_{0.5}$ $(2r + 1) / (2T)$ **(6-56)**

where $X^2_{1-\gamma}(n)$ is the $100\ (1 - \gamma)$th percentile of a chi-square distribution with n degrees of freedom.

Posterior variance: $(2r + 1) / (2 T^2)$

Posterior 100 (1 - γ)% symmetric probability interval:

$$\left(X^2_{\gamma/2}(2r + 1) / (2T), X^2_{1-\gamma/2} (2r + 1) / (2T) \right) \qquad (6\text{-}57)$$

6.3.5.3 Gamma Prior Distribution

The prior density is

$$\frac{\beta_o^{\alpha_o}}{\Gamma(\alpha o)} \lambda^{\alpha_o - 1} \exp(-\beta_o \lambda)(\lambda > 0) \qquad (6\text{-}58)$$

where the positive shape parameter α_0 can be interpreted as the prior number of failures in β_0 prior total operating time. (β_0 , also positive, is the scale parameter.)

The other formulas are:

Prior mean: α_0 / β_0

Prior median: $X^2_{0.5} \ (2\alpha_0) / 2\beta_0$ (6-59)

Prior variance: α_0 / β_0^2

Prior 100(1 - γ)% symmetric probability interval:

$$\left(X^2_{\gamma/2}(2\alpha_0) / (2\beta_0), X^2_{\gamma/2} (2\alpha_0) / (2\beta_0) \right) \qquad (6\text{-}60)$$

Posterior density:

$$\frac{(\beta_0 + T)^{\alpha_0 + r} \lambda^{\alpha_0 + r - 1} \exp[-(\beta_0 + T)\lambda] \quad (\lambda > 0)}{\Gamma(\alpha_0 + r)} \qquad (6\text{–}61)$$

Posterior mean: $(\alpha_0+r) / (\beta_0 + T)$

Posterior median: $X^2_{0.5} (2\alpha_0+2r) / (2\beta_0 + 2T)$ **(6-62)**

Posterior variance: $(\alpha_0+r) / (\beta_0 + T)^2$

Posterior 100(1 - γ)% symmetric probability interval:

$$\left(X^2_{\gamma/2}(2\,\alpha_0 +2r) / (2\beta_0 +2T), X^2_{1\text{-}\gamma/2}\ (2\,\alpha_0\ +2r) / (2\beta_0 +2T\right) \qquad \textbf{(6-63)}$$

See also the other cases for a gamma prior discussed in Section 6.3.3.8. Table 6-2 summarizes the main features of all of the cases.

6.3.5.4 Log-normal Prior Distribution

The prior density is

$$\frac{1}{\sigma\,\lambda\,\sqrt{2\pi}}\ \exp\left[- (\ln\lambda - \xi)^2 / 2\,\sigma^2 \right]\quad (\lambda <0) \qquad (6\text{–}64)$$

The prior moments, etc., are given in Section 6.3.3.9, and the prior 100(1 - γ)% symmetric probability interval is:

$$\left(\exp (\xi - Z_{1\text{-}\gamma/2}\sigma), \exp (\xi + Z_{1-\gamma/2}\sigma) \right) \qquad \textbf{(6-65)}$$

As before, the posterior distribution cannot be obtained in closed form. However, the discrete approximation in Equation 6-10 can be used to approximate the posterior distribution, or numerical integration can be used in conjunction with Equation 6-9. Then f(λ_i) denotes the area under the log-normal prior in the vicinity of λ_i and L(E | λ_i) denotes the likelihood (density function) above evaluated at the chosen discrete set of values λ_i (i = 1,2,...,m).

Case 1. Two prior percentiles λ_{p_1}, λ_{p_2} estimated.

Prior model for each component failure rate	Posterior distribution after observe (r, t), r ≥ o , t > o
Gamma $f(\lambda) = g(\alpha_0, \beta_0)$, with α_0, β_0 found from Martz and Waller (1982) curves. If want convenient posterior distribution and/or percentiles, approximate α_0 by next higher multiple of 1/2, so that $\lambda = \frac{1}{2\beta_0} X^2 (2\alpha_0)$ $\overline{\lambda} = \alpha_0 / \beta_0$ $\lambda_p = \frac{1}{2\beta_0} X^2_p (2\alpha_0)$	$f(\lambda) = g(\alpha, \beta)$ $\overline{\lambda} = \alpha / \beta$ or, if prior α approximated by next higher multiple of 1/2, $\lambda = \frac{1}{2\beta} X^2 (2\alpha)$ $\overline{\lambda} = \alpha / \beta$ $\lambda_p = \frac{1}{2\beta} X^2_p (2\alpha)$

Case 2. Prior mean $\overline{\lambda}$ and one percentile λ_p estimated *

Prior model for each component failure rate	Posterior distribution after observe (r, t), r ≥ o , t > o
X^2 - gamma (after approximating α by next higher multiple of 1/2) $\lambda = \frac{1}{2\beta_0} X^2 (2\alpha_0)$ Find $\alpha_0 = \nu/2$ from $X^2_p(\nu)/\nu = \lambda_p/\overline{\lambda}$ (since X^2 mean = ν)	$\lambda = \frac{1}{2\beta} X^2 (2\alpha)$ $\overline{\lambda} = \alpha / \beta$ $\lambda_p = \frac{1}{2\beta} X^2_p (2\alpha)$

Case 3. Prior mean $\overline{\lambda}$ or one percentile λ_p estimated

Prior model for each component failure rate	Posterior distribution after observe (r, t), r ≥ o , t > o
Exponential - Gamma ($\alpha_0 = 1$) $\lambda = \frac{1}{2\beta_0} X^2 (2)$	$\lambda = \frac{1}{2\beta_0} X^2 (2+2r)$ $\overline{\lambda} = (1+r)/\beta$ $\lambda_p = \frac{1}{2\beta} X^2_p (2+2r)$

Case 4. No prior information

Prior model for each component failure rate	Posterior distribution after observe (r, t), r ≥ o , t > o
Noninformative, conceptualized by $g(\alpha_0, \beta_0)$ with $\alpha_0 = 1/2$ $\beta_0 = 0$	$\lambda = \frac{1}{2t} X^2 (1+2r)$ $\overline{\lambda} = (1+r)/(2t)$ $\lambda_p = \frac{1}{2t} X^2_p (1+2r)$

Notes: $g(\alpha, \beta)$ = gamma distribution with parameters α, β

t = Length of observation time

r = Number of observed failures during t

$\alpha = \alpha_0 + r$

$\beta = \beta_0 + t$

$\overline{\lambda}$ = Mean failure rate

λ_p = P_{th} percentile of failure rate uncertainty distribution

$X^2_p(\nu)$ = P_{th} percentile of X^2 distribution with ν degrees of freedom

* See caution note in text

TABLE 6-2. SUMMARY OF APPLICABLE FAILURE RATE INFERENCE METHODS EMPLOYING THE GAMMA DISTRIBUTION (HEUBACH AND PHILIPSON, 1985)

6.3.6 Example: Failure of Diesel Generators To Start

Presented below is an example from Apostolakis et al.(1980). The frequency with which diesel generators fail to start (measured in terms of the failure rate per demand) was assumed in the Reactor Safety Study (Nuclear Regulatory Commission, 1975) to have a log-normal distribution with 5th and 95th percentiles of 10^{-2} and 10^{-1}, respectively. Using the procedure outlined in Section 6.3.3.9, it is found that $\xi = 3.45$ and $\sigma = 0.70$ are the two log-normal parameter values. The prior mean, mode, median, and variance are then 0.04, 1.9×10^{-2}, 3.2×10^{-2}, and 1×10^{-3}, respectively.

Suppose now that the evidence E from a certain facility consists of $r = 5$ failures in $n = 227$ test demands. Table 6-3 shows the discretized log-normal prior and calculations required to compute the corresponding posterior distribution by means of Equation 6-2; values smaller than 10^{-4} have been treated as equal to zero.

Figure 6-3 shows a plot of the discretized prior and posterior distributions and gives a graphic illustration of the change in the generic prior brought about by the influence of the facility-specific evidence. The posterior mean and variance are computed to be 0.025 and 8.2×10^{-5}, respectively. The effects of the facility-specific evidence are, first, to shift the distribution of the failure-to-start probability toward lower values and, second, to reduce the dispersion.

Another alternative Bayesian procedure is to approximate the binomial likelihood with a Poisson distribution and to assign a conjugate gamma prior distribution to the corresponding failure rate. Taking the 5th and 95th percentiles to be 10^{-2} and 10^{-1}, respectively, and using the procedure of Martz and Waller (1982) yields a gamma prior distribution with the shape parameter $\alpha_0 = 2.4$ and the scale parameter $\beta_0 = 52.68$. Using the results in Section 6.3.5.3, the posterior distribution is another gamma distribution with the shape parameter 7.4 and the scale parameter 279.68. The corresponding posterior mean and variance are computed to be 0.026 and 9.5×10^{-5}, respectively. The posterior 5th, 50th, and 95th percentiles are also easily computed to be 0.013, 0.038, and 0.045, respectively.

Failure rate (failure to start)	Prior probability	Likelihood	(Prior) x (Likelihood)	Posterior probability
.0087	.0500	.0343	.0017	.0206
.0115	.0587	.0750	.0044	.0529
.0154	.0967	.1320	.0128	.1535
.0205	.1350	.1734	.0234	.2815
.0274	.1596	.1544	.0246	.2963
.0365	.1596	.0820	.0131	.1572
.0487	.1350	.0218	.0029	.0353
.0649	.0967	.0023	.0002	.0027
.0866	.0587	.0001	.0000	.0000
.1155	.0500	.0000	.0000	.0000
Sum	1.0000	.0831	1.0000	

[a] From Apostolakis et al. (1980).

Table 6-3. Estimation of Diesel Generator Failure to Start by the Bayesian Method[a]

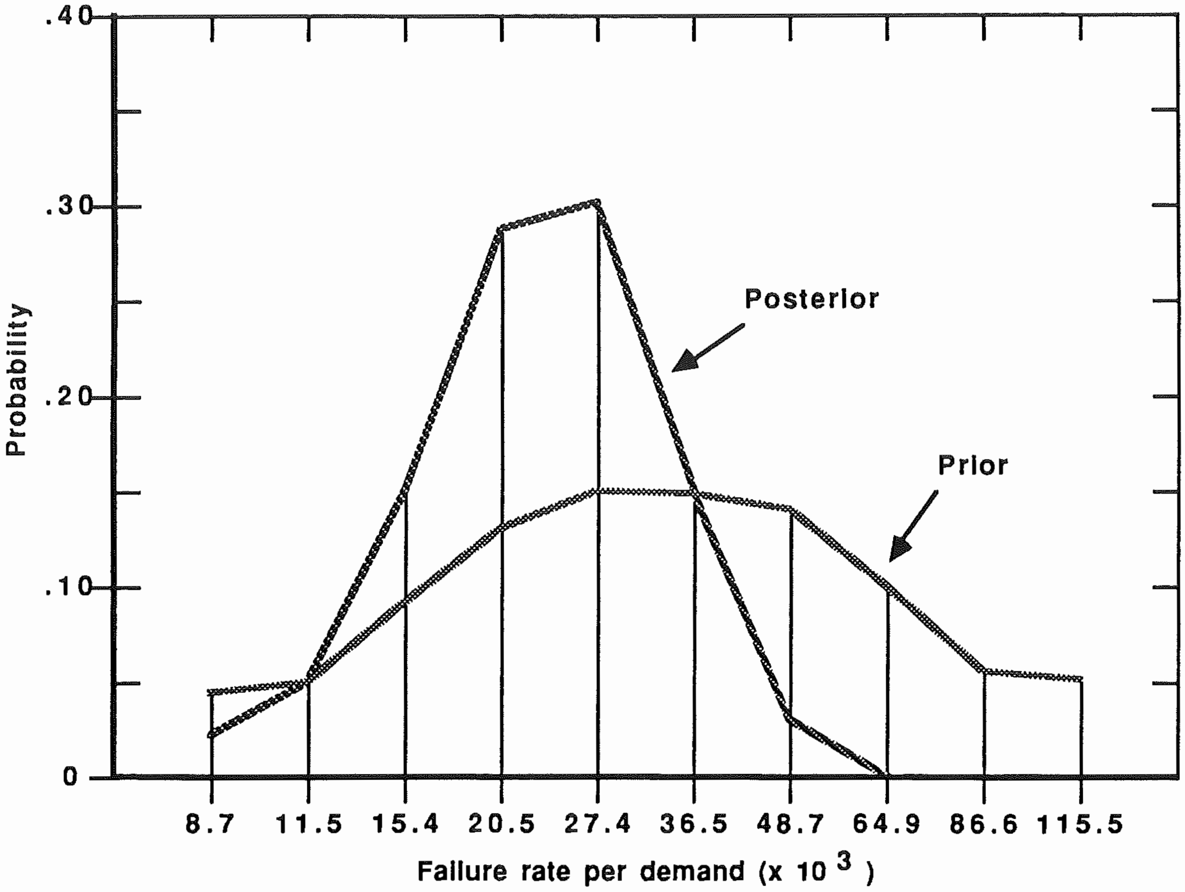


FIGURE 6-3. PRIOR AND POSTERIOR HISTOGRAMS FOR DIESEL GENERATOR FAILURE TO START (FROM APOSTOLAKIS ET AL., 1980)

Consider now the estimation of the probability of diesel generator failure to start by the classical methods of Section 6.2.1. The data, f/n = 5/227, lead to a maximum-likelihood estimate of p = .022, with a standard deviation of .0097. Note that the square of this standard deviation is 9.4×10^{-5}, which is about the same as the above Bayesian posterior variance.

Table 6-4 gives lower and upper classical confidence limits on the failure-to- start probability for a variety of confidence levels. It presents both the exact evaluations noted in Section 6.2.1 and the chi-square approximations. Both sets of confidence limits are shown to four decimals only to illustrate the close agreement between the exact and the approximate bounds for these data.

Because of the discretizing that is used, it is difficult to compare the Bayesian results in Table 6-3 with the classical results in Table 6-4. Qualitatively, however, both analyses suggest strongly that the failure probability of interest is between .01 and .05. As one method of comparison, note that data of 7.5 failures in 300 demands would yield a maximum likelihood estimate and a squared standard deviation essentially equal to the Bayesian posterior mean and variance; thus, the Bayesian prior effectively contributed additional data of 2.5/73 to the results.

In general, the different analyses of these data agree quite closely, even though the interpretation varies considerably. The main reason for this agreement in this example is the rather large quantity of facility-specific data, which results in a likelihood that dominates the prior distribution in the Bayesian analysis and so diminishes the impact of the Bayesian-particular aspects.

6.4 DERIVATION OF SYSTEM-LEVEL INFERENCES

Bayesian component-level data developments facilitate system-level inferences with their incorporation of explicit uncertainty distribution functions. No concern arises as in classical procedures on the proper use of component-level confidence bounds in deriving system-level confidence bounds. In principle, given the uncertainty distributions for all components' failure rates, the uncertainty distribution at the system-level can be established by Monte Carlo procedures for any series-parallel system (including systems with dependent component failures if the associated conditional failure probabilities are estimated). In some cases, analytical procedures can be applied and usually good,

Table 6-4. Classical Confidence Limits on the Probability of Diesel Generator Failure to Start (Five Failures in 227 Attempts)[a]

	Exact solution		Chi-square approximation	
Confidence level (%)	Lower	Upper	Lower	Upper
50	.0205	.0249	.0206	.0249
75	.0149	.0325	.0148	.0327
90	.0108	.0405	.0107	.0407
95	.0087	.0458	.0087	.0463
97.5	.0072	.0507	.0072	.0513
99	.0057	.0567	.0056	.0577

[a] From Apostolakis et al. (1980).

approximate system-level results obtained relatively simply. The case of a series system with independent component failures and gamma component-level uncertainty distributions for the failure rates is given here as a particular example. (See also Philipson and Tran, 1984, and Heubach and Philipson, 1985, 1986, in which series properties including the additivity of failure rates are associated with sets of failure modes occurring during successive launches of a space or a missile vehicle.) A further approximation may sometimes also enable the application of the simple analytical procedures to the series-parallel case, as well. This will also briefly be indicated.

6.4.1 Series System

The basic step is to substitute for each component's prior uncertainty gamma distribution a weighted X^2, as discussed in Section 6.3.3.8. When the gamma distribution's α_0 value is a multiple of 1/2, this substitution is exact, producing an expression for the variate λ ,

$$\lambda = \frac{1}{2\beta_0} X^2 (2\alpha_0) \qquad \textbf{(6-66)}$$

If α_0 is not a multiple of 1/2, this result is obtained approximately and conservatively by rounding up α_0 to the next higher multiple of 1/2. The approximation should be satisfactory except, perhaps, when $\alpha_0 <1$. As seen in Figure 6-4, when $\alpha_0 <1$ the gamma density has an exponential shape, with no mode and monotonically decreasing. See Heubel and Philipson (1986) for an exploration of the accuracy of this approximation.

The prior distribution for each component's λ is next revised with any available sample data (r, t), $r \geq 0$, $t > 0$ producing a new weighted X^2 variate,

$$\lambda = \frac{1}{2\beta} X^2 (2\alpha) \qquad \textbf{(6-67)}$$

with

$$\alpha = \alpha_0 + r$$
$$\beta = \beta_0 + t$$

as discussed before.

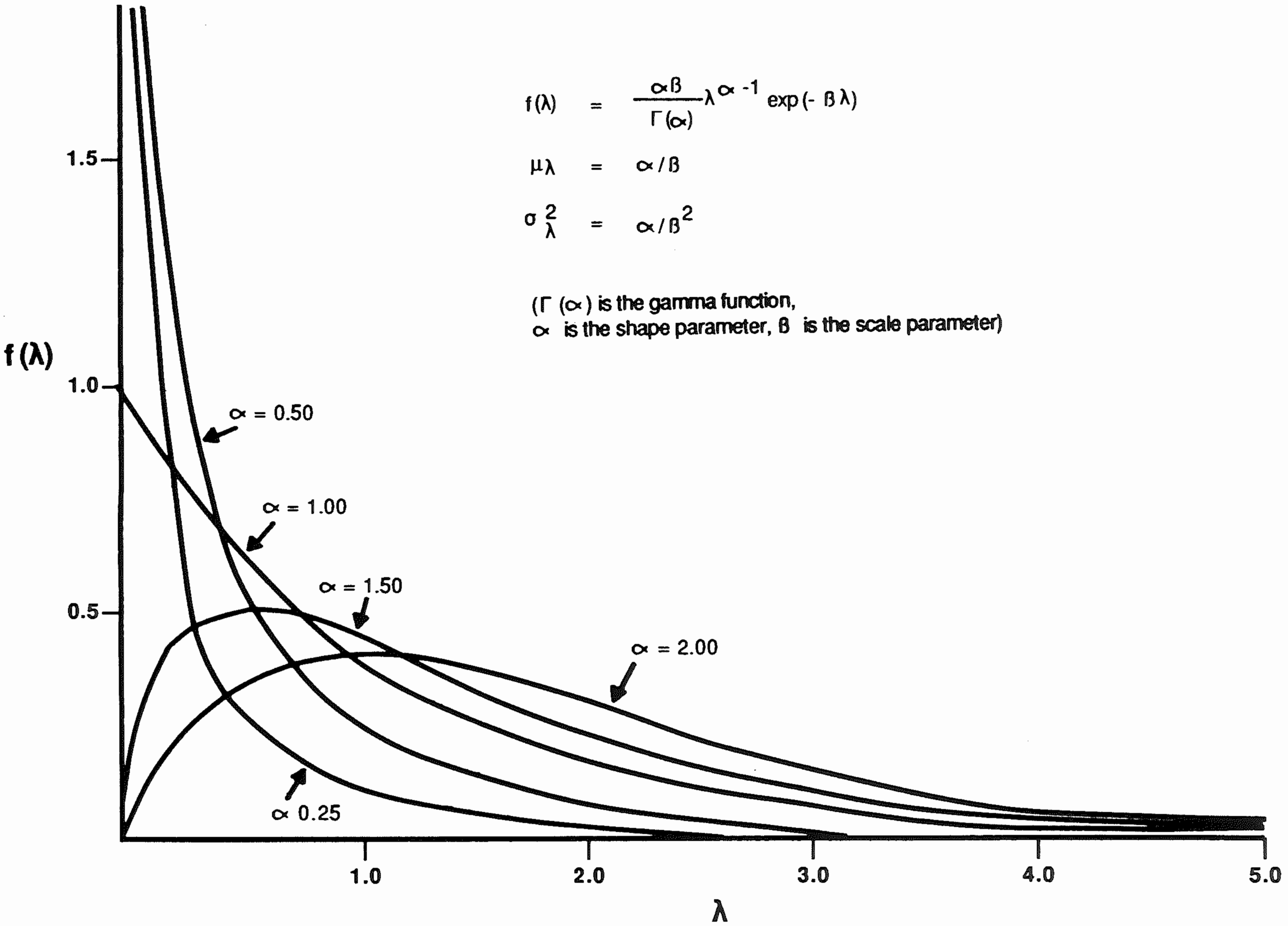


FIGURE 6-4. GAMMA DISTRIBUTION FAMILY (WITH $\beta = 1.0$)

Then, as shown in Mann, Schafer, and Singpurwalla (1974), the failure rate of a system made up of a series of such components i with failure rates $\lambda_{i,}$ which is thus a sum of weighted X^2 variates, is also itself a weighted X^2 variate, with mean and variance

$$m = \sum_i (\alpha_i / \beta_i)$$

$$v = \sum_i (\alpha_i / \beta_i^2) \tag{6-68}$$

Using the very accurate Wilson-Hilferty approximation for a X^2 percentile (Mann, Schaefer, and Singpurwalla, 1974), this gives for the system-level failure rate's point estimate and pth percentiles,

$$\overline{\lambda} = m$$

$$\lambda_p = m \left(1 - \frac{v}{9m^2} + Z_p \frac{\sqrt{v}}{3m}\right)^3 \tag{6-69}$$

where Z_p is the pth percentile of the standard normal deviate.

6.4.2 Series-Parallel System

These same simplified procedures can also be applied to series-parallel systems (with all components on-line) if a further approximation is made. Whether or not this approximation is satisfactory needs to be investigated case-by-case. However, it can always be assured to be conservative, in the sense that the approximate value of the system-level failure rate developed will be no larger than the non-approximated value of the instantaneous failure rate at any time during the system's period of operation. This also implies that the reliability to any given time computed with the value of the approximate failure rate will be no greater than the reliability computed with the value of the instantaneous failure rate at the given time.

The idea is simply to reduce the series-parallel system to an approximating system of units in series by obtaining an "effective constant failure rate" (ECFR) for each unit with parallel components. If the parallelisms thereby approximated away are low in the system's network diagram this procedure can be satisfactorily accurate. If the system has parallel components near the top of its network diagram, the approximation's accuracy may not be satisfactory. However, when it is satisfactory, the procedures of Section 6.4.1 then again apply.

The derivation of the ECFR for a unit containing parallel components to enable this is as follows. It is assumed that all the individual components have constant failure rates during the system's operating period, $0 \leq t \leq T$. Then it can be shown that the instantaneous failure rate λ (t) of the unit is non-decreasing and has a maximum value at t=T. Compute by standard methods the reliability R(T) of the unit to time T. Then λ (T) is found as

$$\lambda(T) = \frac{-\ln R(T)}{T} \qquad \textbf{(6-70)}$$

since

$$R(t) = e^{-\lambda(t)t} \qquad \textbf{(6-71)}$$

by definition of the instantaneous failure rate (or, as it is also termed, the "hazard function"). If λ (T) is used as an approximation to $\lambda(t)$ in $0 \leq t \leq T$, then since $\lambda(T) \geq \lambda(t)$, $R(T) \leq R(t)$ throughout this time period and thus the approximation is conservative.

The approximate <u>system</u> reliability to time T that the use of each series unit's $\lambda(T)$ implies must also be no greater than the true system reliability, and so the approximate system failure rate (a constant since it derives from a series of units with constant actual or effective failure rates) must be at least as large as its true instantaneous failure rate throughout $0 \leq t \leq T$.

The foregoing derivation provides only a point estimate for the ECFR or the reliability for each unit and for the system. To obtain an approximate uncertainty distribution for the ECFR, assume the point estimate is the median of the uncertainty distribution, obtain another percentile, and fit a gamma distribution as discussed previously. (The obtaining of the second percentile is done most simply from generic data or judgment applied directly to the unit; or otherwise by Monte Carloing the combination of the gamma uncertainty distributions of the parallel components making up the unit.) The approximate system failure rate's uncertainty distribution is then established as in the first part of this section, and its percentiles are easily computed as before.

6.5 DEPENDENT FAILURES DATA

Events are dependent (or, perhaps better, non-independent) when their probabilities of occurrence are all affected by the occurrence of some particular event. Thus, common cause failures are events in which several subsystems or components fail (their probabilities of failure all become 1.0) if a common cause event occurs (e.g., a structural failure causes all control cables, supposed to provide redundancy, to break simultaneously). In a second case, the several components' probabilities of failure all increase (not necessarily to 1.0) because of a common cause event (e.g., a temperature control device fails and allows all components to overheat; or, inadequate maintenance degrades all components). These are two basic kinds of dependent failures. Variations are noted in Section 3.7 of Nuclear Regulatory Commission (1983).

Mathematically, in general, events are dependent if the probability of their joint occurrence is not equal to the product of their individual marginal probabilities of occurrence, but requires incorporation of their conditional probabilities of occurrence that depend on whether or not other events occur.

In the simplest case of two dependent events, A and B:

$$p\,(A \text{ and } B) = p\,(A|B)\,p(B) \neq p\,(A)p(B) \qquad \textbf{(6-72)}$$

The general logic tree formulation and analysis of dependent failures was discussed in Section 4. In this section, the treatment of a certain special class of dependent failures and the development of data for this treatment are described. This class is that of multiple dependent failures occurring with frequently used sets of equipments intended to be redundant (e.g., sets of diesel generators of a given type used at many nuclear power plants for auxiliary electric power generation), for which a statistically significant sample of multiple failures has been observed. If the sample is adequate, and undue complexity in a fault tree does not arise when all the possible multiple failures are included individually, straightforward inferences of the probabilities of occurrence of the possible multiple failures can be employed. Otherwise, certain models have been developed for multiple failures which, in particular, enable simplifications in the incorporation of these failures in fault trees. However, while a discussion of these models and their data development is included here for completeness, it is not considered that such models will be usable or their data obtainable very often in NASA applications, with their generally one- or few-of-a-kind characteristics.

6.5.1 The Beta-Factor Model

This model, developed by Fleming et al. (1975), assumes that the total failure rate of each unit U in a multiple redundant set can be expressed as the simple sum,

$$\lambda_U = \lambda_{UI} + \lambda_{UC} \qquad (6\text{-}73)$$

where λ_{UI} is the unit's independent failure rate; λ_{UC} is the additional dependent failure rate. One then defines the parameter

$$\beta_U = \frac{\lambda_{UC}}{\lambda_U} = \frac{\lambda_{UC}}{\lambda_{UI} + \lambda_{UC}} \qquad (6\text{-}74)$$

the proportion of the total failure rate of unit U due to dependent failures.

It can then be shown, for example, that the probability of failure of a redundant system that is operable if at least one of two identical units (with the same λ_{UI} and β_U) is operable is

$$q_s = (1 - \beta_U)^2 \lambda^2{}_{UI} (1-\beta_U \lambda_{UI}) + \beta_U \lambda_{UI} \qquad (6\text{-}75)$$

q_s is estimated by estimating λ_{UI} and β_U from experience data. The use of q_s in a fault tree enables consideration of just the single event -- the system fails -- rather than the distinct events corresponding to both units failing from (a) independent causes and (b) from common causes. For more complex redundant systems this simplification becomes significant. However, in such systems, the beta-factor model does not distinguish between different numbers of multiple failures but must perhaps over-conservatively assume <u>all</u> units necessarily fail when a multiple failure occurs. The binomial failure rate model next discussed improves on this.

6.5.2 Binomial Failure Rate (BFR) Model

This model is a specialization of a more general one established by Marshall and Olkin (1967). See Vesely (1977) for details beyond those given here. Also see Atwood (1983) for additional discussion of data development procedures in applications of the model.

The basic concept of the BFR model has much in common with that of the beta-factor model, but it is more correct statistically, and it adds an ability to distinguish and employ data on partial failures of a system consisting of a redundant set of components. Table 6-5 from Atwood (1983) lists the model's input and output quantities of interest. It exhibits that the common cause event is a "shock," occurring at a given rate, which may be lethal with probability 1.0 to all components in the set, or may be nonlethal and only increase the rate of occurrence of common cause failures of specified components and numbers of components in the set. The total failure rate for a given component is then the sum of its independent failure rate, its lethal failure rate (which is the same as the rate of occurrence of a lethal shock), and its nonlethal shock-caused failure rate. These rates must be the same for all components in the set. The several necessary parameters in the estimates of these rates are themselves estimated from observational data on single and multiple failures of sets of components subject to common cause shocks and on the occurrences of the shocks (see Atwood, 1983). Multiple-component system failure rates are then developed with combinatorial formulas, as illustrated in Table 6-5.

6.5.3 Comparison of the Data Requirements of the Beta-Factor and BFR Models

The beta-factor model employs estimates developed from observational data of λ , the total failure rate, and β the common cause-induced portion of the total failure rate, or, equivalently, the independent and common cause failure rates, λ_I and λ_C . The BFR model requires estimates of the shock occurrence rate and the probability of failure of each component given that a shock occurs. Thus, in addition to data on the usual independent component failure rate, the beta-factor model requires data on the rate of common cause failures of a set of components sufficient to estimate one parameter, β. The BFR model requires data to estimate two additional parameters: the rate of occurrence of shocks and the conditional rate of failure of the component given that the shock occurs (taking into account whether all components in the set fail when the shock occurs or only a fraction of them). Both models assume the data are associated with identical components. The beta-factor model is somewhat less sensitive to departures from this assumption.

TABLE 6-5. QUANTITIES OF INTEREST (ATWOOD, 1983)

λ	=	failure rate for an individual component, not counting failures due to common cause shocks
μ	=	rate of nonlethal shock occurrences
p	=	probability that a specific component fails, given that a nonlethal shock occurs
$\lambda_+ = \mu(1 - q^m)$	=	rate of nonlethal shocks that cause at least one component failure, i.e., rate of visible nonlethal shocks (here, q = 1-p)
ω	=	rate of lethal shock occurrences
$r_1 = \lambda + \mu p + \omega$	=	rate at which a specific component fails, either due to individual failure or due to a shock
$r_k = \mu p^k + \omega$, for $k \geq 2$	=	rate at which a specific set of k components fails simultaneously (due to a shock)
$r_{1/m} = m\lambda + \lambda_+ + \omega$	=	rate at which at least 1 out of m components fails
$r_{k/m} = \mu \sum_{i=k}^{m} \binom{m}{i} p^i q^{m-i} + \omega$ for $k \geq 2$	=	rate at which at least k out of m components fail simultaneously
$\beta = [\mu p(1 - q^{m-1}) + \omega]/ r_1$	=	long-term fraction of component failures that occur in multiple failures

Clearly, extensive observational data samples are necessary to estimate these parameters with good confidence. But, in addition, it must be recognized that not many real systems, especially in NASA, can exhibit even approximately the simple commonalties required for the models to apply with any fidelity. The two models attempt to subsume complex properties in simple formulas. They may well be inadequate in particular cases. Care must be taken in using them even when they appear to have some applicability.

6.6 HUMAN ERROR DATA

A brief summary of the domain of human reliability modeling and human error probability (HEP) data development is first given. It exhibits the basically different approaches that have been developed for assessing mechanistic errors ("slips") in carrying out routine, often prescribed, procedures; versus errors ("mistakes") in the cognitive processes of interpretation and decision-making especially pertinent in contingency actions. Following the summary, descriptions are provided of certain of the major techniques for developing HEP data that can quantify basic events in fault trees incorporating models of human operation or maintenance activities. It must be emphasized, however, that human reliability analysis is a complex and still evolving process, and original sources should be referred to for completeness and fuller understanding.

6.6.1 Summary

Human error types can be classified according to various factors. Figure 6-5 indicates that twelve categories of errors can result.

Mistakes are not usually modeled in normal task activities (category 3). The driving assumption is that each task plan is specified by management and detailed in technical specifications, so that the decision elements of the task are, thus, highly reliable. Unspecifiable tasks obviously cannot be modeled, although for some applications enough data may exist to estimate an overall failure rate for this category (category 4). Commission failures that are due to mistakes are often not modeled (categories 8 and 12). This is because it is assumed that other error types dominate their occurrence probabilities and the presence of extensive instrumentation makes their effects detectable and able to be mitigated.

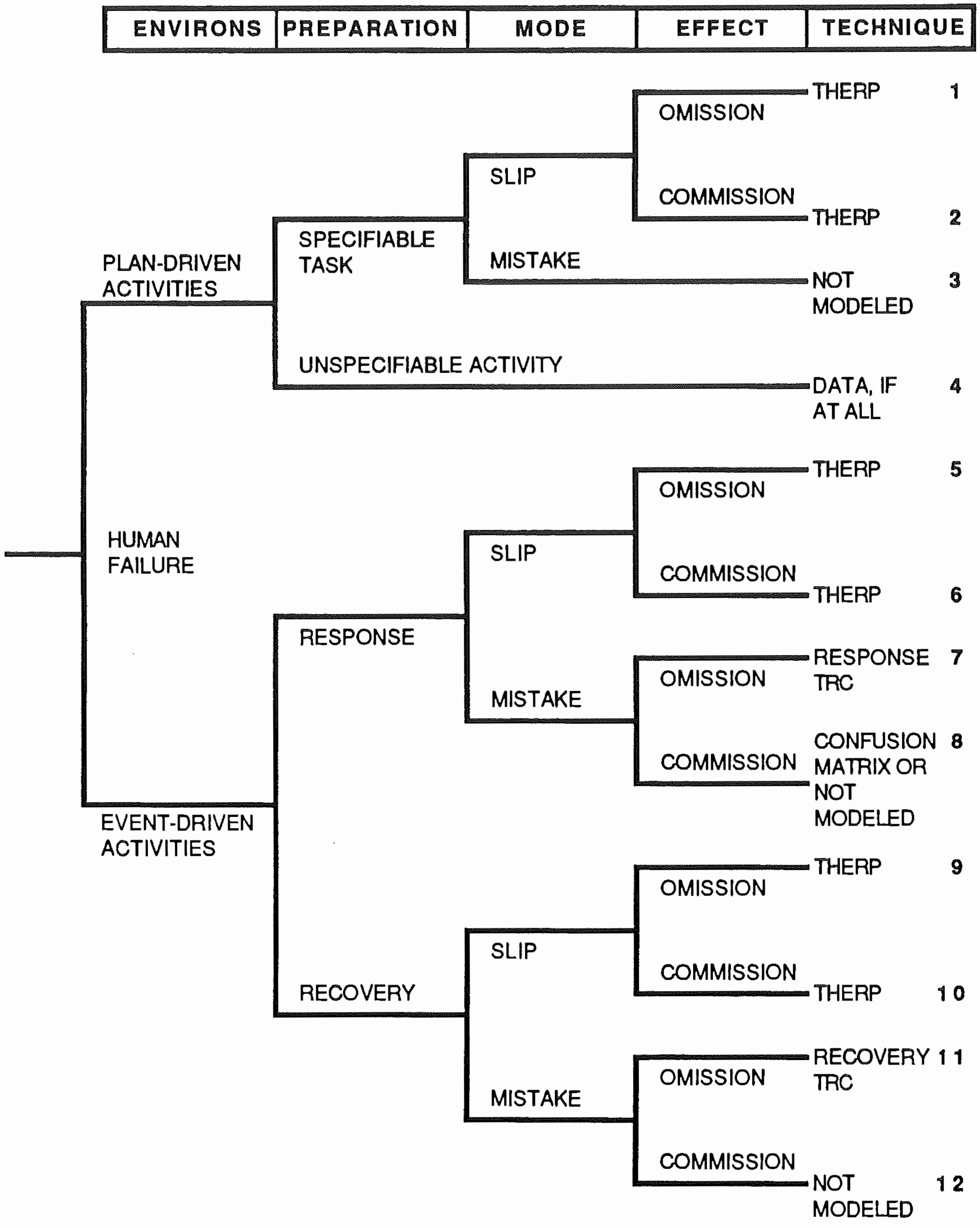


FIGURE 6-5. CLASSIFICATION SYSTEM FOR HUMAN ERRORS
(ADAPTED FROM DOUGHERTY AND FRAGOLA, 1988)

In Figure 6-5, THERP refers to the Technique for Human Error Rate Prediction (Swain and Guttmann, 1983) which is a now standard procedure for estimating human error ("slip") probabilities (HEPs) in routine, often prescribed tasks. For event-driven activities, such as contingency actions, in which mistakes are likely to be more significant than slips, THERP is not applicable, but some techniques that have been developed recently are. Time Reliability Correlations (TRCs) (Dougherty and Fragola, 1988), and the related Operator Action Tree (OAT) (Hall, Fragola, and Wreathall, 1982), are a principal example.

The THERP and OAT/TRC approaches for developing HEPs for slips and mistakes, respectively, are discussed in the following two subsections.

6.6.2 THERP

This discussion draws largely on the Human Reliability Analysis (HRA) chapter in the PRA Procedures Guide (Nuclear Regulatory Commission, 1983), based on Swain and Guttman (1983), which should be referred to for further details.

Most of the available estimates of human error probabilities represent extrapolations from human-error data based on tasks performed in various contexts which are behaviorally similar to those of interest, especially those performed in nuclear power plants. The tasks are behaviorally similar because they may involve the same types of cues, interpretations, response requirements, and responsibilities as those performed in nuclear power plants. However, in those cases for which an analyst can find better human performance data, he should use them.

Nearly all quantified human error probabilities relate to routine human actions. For some operations, cognitive errors are critical (e.g., errors in evaluating display indications). There has been very little information on errors of interpretation or decisionmaking (i.e., errors in the thought process) until recently. Such errors are not considered here but are discussed in the following subsection.

Nominal values for the probabilities of given human actions as well as uncertainty bounds have been developed (Swain and Guttmann, 1983). The nominal values reflect the best estimate (based on available data and on judgment) of the probability of a particular error in a generic sense. The uncertainty bounds are considered to approximate the middle 90-percent range of the human error probabilities to be expected under all possible scenarios for a particular action. These uncertainty bounds are based on subjective judgment rather than on actuarial data and are not meant to represent statistical confidence limits.

There are several sources of uncertainty in the generic HEP values. The variability of human performance is reflected in the differences among operating or maintenance personnel--differences in skill, experience, and other personal characteristics. There can be wide variations in specific environmental situations including man-machine interfaces, and in other physical aspects of the tasks to be performed or in the response requirements under which the operator must act. Only some of this variation in such "performance-shaping factors" has been accounted for in the available data by providing different estimates of human error probabilities for different sets of influencing factors. The width of the uncertainty bounds surrounding each estimated nominal probability represents an attempt to account for the residual uncertainty.

6.6.2.1 Assumptions

Almost all of the HEP estimates that have been developed are based on a set of common assumptions that limit or restrict the use of the data:

1. The system or other application is operating under normal conditions. There is no emergency or other state that would produce in the operators a level of stress other than the optimal.

2. In performing the operations, the operator does not need to wear protective clothing.

3. A generally accepted level of administrative control is in effect.

4. The tasks are performed by licensed, qualified personnel, such as operators, maintainers, or technicians. They are assumed to be experienced--to have functioned in their present positions for at least six months.

5. The environment for the activity is not adverse. The levels of illumination and sound and the provisions for physical comfort are at least adequate.

6.6.2.2 HRA Procedures Employing THERP

Given that the foregoing general assumptions hold, a human error analysis (HRA) is performed with the steps indicated in Figure 6-6 to arrive at the event probabilities required for a fault tree.

- Review of System Characteristics and of Information from System Analysts

  For a given scenario or sequence of events, the system analysts identify human actions that directly affect the system-critical components. With the light of the information obtained from the system review, the human-reliability analyst assesses these actions in the context of their actual performance; the objective is to determine whether these actions can be affected by factors that may have been overlooked by the system analysts. For example, if performance on a non-critical element subsequently affects performance on a system-critical element, this effect must be considered, even though that task in itself is not important to the reliability of the system as defined by the system analysts.

- Talk-Through of Procedures

  Sometimes performed in conjunction with the system survey and sometimes at a later date during interviews with personnel, talk-throughs of the procedures in question are an important part of any human reliability analysis. They are conducted by the human reliability analyst and performed by system operations or maintenance personnel. Performance specifics are identified along with any time requirements, personnel assignments, skill-of-the-craft requirements, alerting cues, and recovery factors.

  The information obtained in a talk-through helps the analyst to account for the effects of a situation's performance-shaping factors. Modifications made to available nominal HEP values are based on this information.

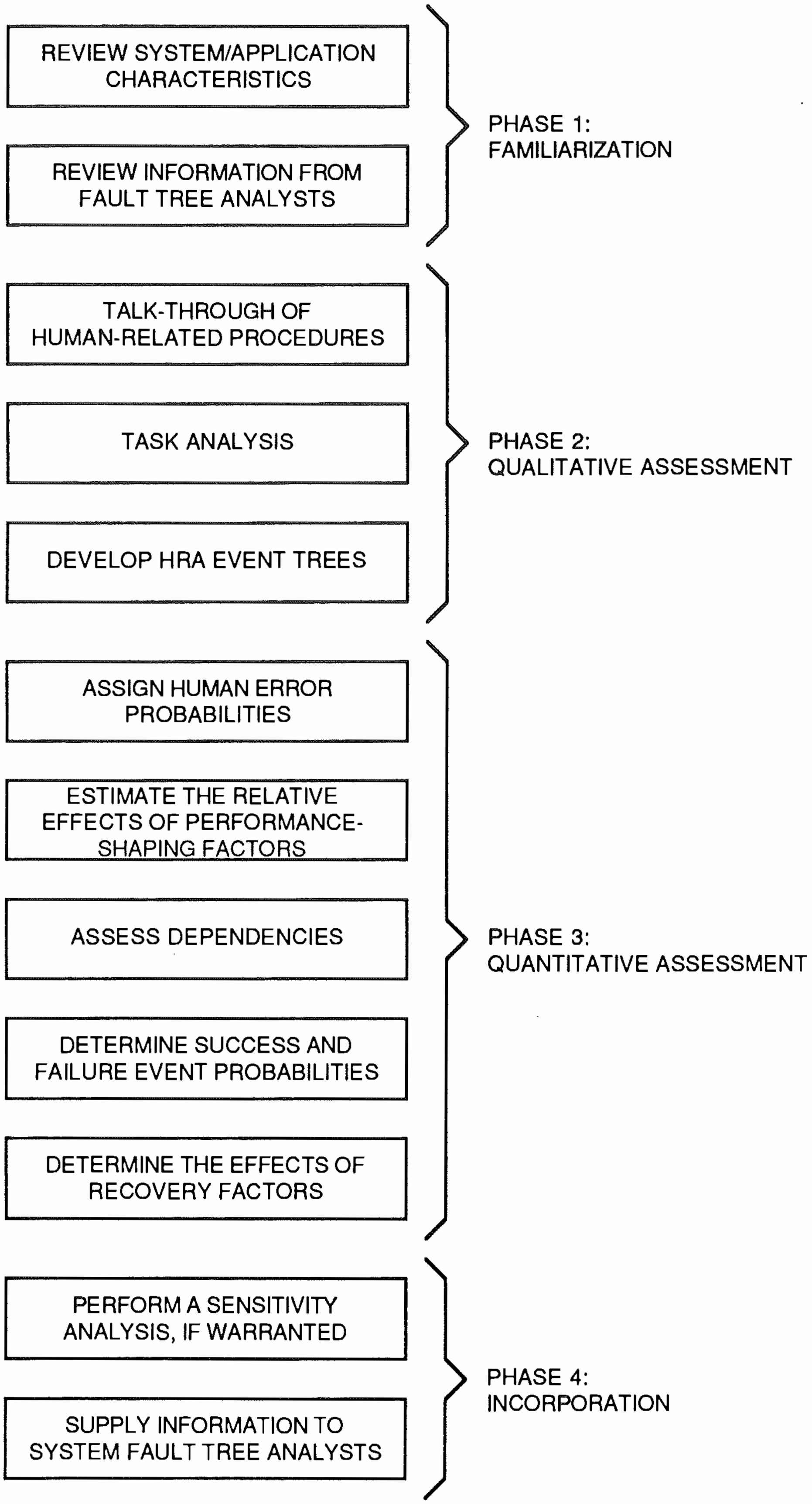

REVIEW SYSTEM/APPLICATION CHARACTERISTICS
REVIEW INFORMATION FROM FAULT TREE ANALYSTS
PHASE 1: FAMILIARIZATION
TALK-THROUGH OF HUMAN-RELATED PROCEDURES
TASK ANALYSIS
DEVELOP HRA EVENT TREES
PHASE 2: QUALITATIVE ASSESSMENT
ASSIGN HUMAN ERROR PROBABILITIES
ESTIMATE THE RELATIVE EFFECTS OF PERFORMANCE-SHAPING FACTORS
ASSESS DEPENDENCIES
DETERMINE SUCCESS AND FAILURE EVENT PROBABILITIES
DETERMINE THE EFFECTS OF RECOVERY FACTORS
PHASE 3: QUANTITATIVE ASSESSMENT
PERFORM A SENSITIVITY ANALYSIS, IF WARRANTED
SUPPLY INFORMATION TO SYSTEM FAULT TREE ANALYSTS
PHASE 4: INCORPORATION

FIGURE 6-6. OVERVIEW OF A HUMAN RELIABILITY ANALYSIS EMPLOYING THERP (ADAPTED FROM NUCLEAR REGULATORY COMMISSION, 1983)

- Task Analysis

At this point, a task analysis should be performed. A "task" is defined as an element of activity or performance that can be treated as a unit either because of its performance characteristics or because the task is required as a whole in the overall activity. Only the tasks that are relevant to the safety of the system are considered. A task analysis involves breaking down each task into individual units of behavior. Usually, this breakdown is done by tabulating information about each specific human action. The analysis and the information it yields can be either qualitative or quantitative.

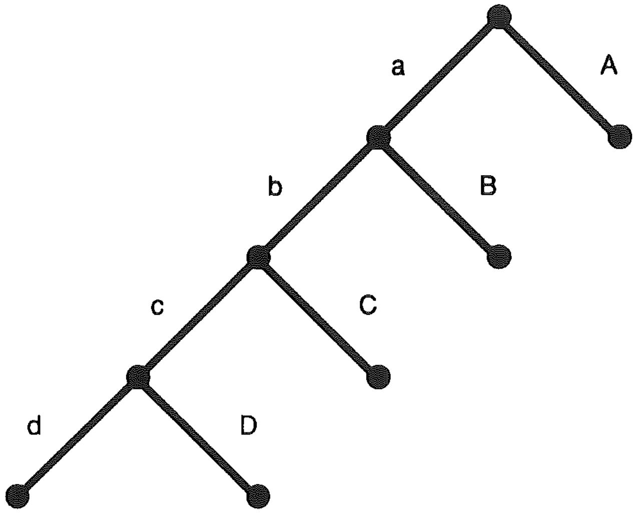
Specific potential errors are identified for each unit of behavior. For every human action appearing in the task-analysis table, likely errors of omission and commission are identified. A human action (or its absence) constitutes an error if it has at least the potential for reducing the probability of some desired event or condition. The existence of this potential should be identified in conjunction with the system analysts.

- Development of HRA Event Trees

Each of the errors defined above is now entered as a binary branch on an HRA event tree. The possible error events should appear on the tree in the order in which they might occur if such order is relevant. The suggested format for HRA event trees is illustrated in Figure 6-7. The product of the HRA event tree is a probabilistic statement of the likelihood of a given sequence of events.

- Assignment of Nominal Human-Error Probabilities

An estimate of the probability of each human error event on the HRA event tree is next derived from the data tables in such sources as the Nuclear Regulatory Commission HRA Handbook (Swain and Guttmann, 1983). (More specifically applicable data can be established if experiments or simulations can be conducted.) Tables of human error probabilities (and the associated uncertainty bounds) for generic task descriptions are available. If there is no exact match between the description of a task in the Handbook and that defined by the task analysis, the



EVENT

A = CONTROL ROOM OPERATOR OMITS ORDERING THE FOLLOWING TASKS

B = OPERATOR OMITS VERIFYING THE POSITION OF MU-13

C = OPERATOR OMITS VERIFYING/ OPENING THE DH VALVES

D = OPERATOR OMITS ISOLATING THE DH PUMP ROOMS

FIGURE 6-7. ILLUSTRATIVE HRA EVENT TREE FOR ACTIONS PERFORMED OUTSIDE A NUCLEAR POWER PLANT CONTROL ROOM (NUCLEAR REGULATORY COMMISSION, 1983)

estimated error probability for a similar task may be able to be used as is, or it may be able to be extrapolated, depending on the degree of similarity between the descriptions. "Similarity" in this context refers to the likeness of required operator behaviors. There can be a high degree of similarity between the performance of two tasks even though the equipment and overall operation are dissimilar.

- Estimating the Relative Effects of Performance-Shaping Factors

  The human error probabilities estimated in the Handbook and other sources for a given task must now be modified to reflect the actual performance situations of concern. For example, if the labeling scheme at a particular facility is very poor, in comparison with those described in Military Standard 1472C (U.S. Department of Defense, 1981) or NUREG-0700 (USNRC, 1981b), the probability of error should be increased toward the upper bound of its uncertainty range. If the tagging control system is particularly good, perhaps the probability for certain errors can be decreased.

  Some of the performance-shaping factors (PSFs) given in the Handbook affect a whole task or a whole procedure, whereas others affect certain types of errors, regardless of the tasks in which they occur. Still other PSFs have an overriding influence on the probabilities of all types of error in all conditions.

- Assessment of Dependencies

  In any given situation, there may be different levels of dependence between an operator's performance on one task and on another because of the characteristics of the tasks themselves or because of the manner in which the operator was cued to perform the tasks. Dependence levels between the performances of two (or more) operators may differ, also. It is essential to keep in mind that the effect of dependence on human error probabilities is always highly situation-specific. The Handbook presents appropriate means for the treatment of such dependencies.

- Estimating Success and Failure Probabilities

  The criteria for system success and failure are supplied by the system analysts. These criteria are used as the basis for labeling the end point of each path through an HRA event tree as a success or a failure. As illustrated in Figures 6-8 and 6-9, probabilities are then assigned to each success or failure event from the estimates in Handbook tables or other sources multiplied by appropriate performance-shaping factors. Then, multiplying the probabilities assigned to each limb in a success or failure path through the HRA event tree provides a set of success and failure probabilities that can then be combined to estimate the total system success and failure probabilities.

- Determining the Effects of Recovery Factors

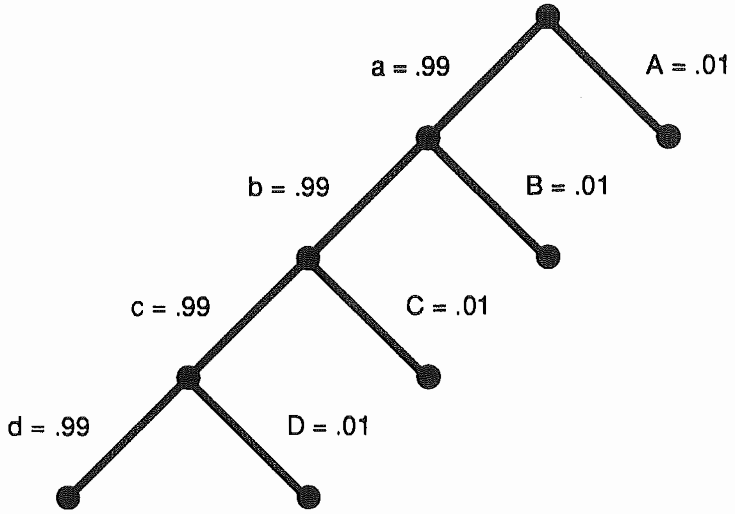
  It is often convenient to postpone consideration of the effects of recovery factors until after the total system success and failure probabilities have been determined. The estimated probabilities for a given task sequence may be sufficiently low without considering the effects of recovery factors so that the sequence does not appear as a potentially significant failure mode. In this case, it can be dropped from further consideration.

- Performing Sensitivity Analysis, If Warranted

  To determine the effect of a single parameter on the total system success or failure probability, a sensitivity analysis can be performed. In this exercise, the value of a given parameter is manipulated and the resulting system success probabilities are compared to judge the impacts of different magnitudes of change.

- Supplying Information to Fault Tree Analysts

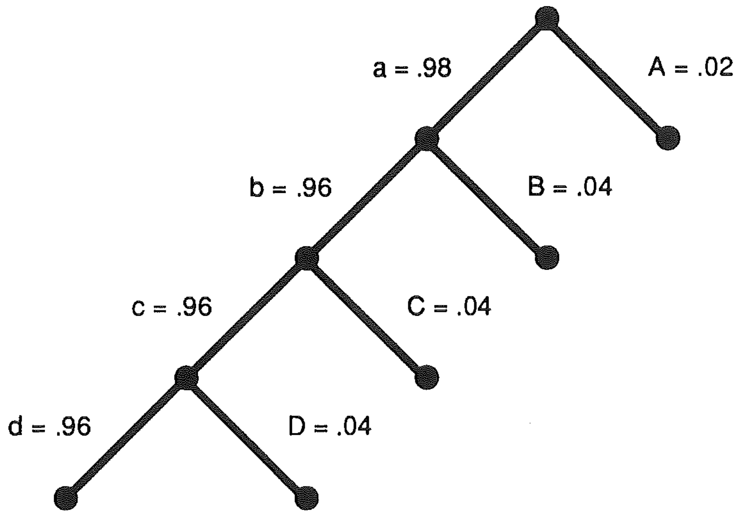
  A copy of each HRA event tree along with a synopsis of the results, a copy of the task-analysis table, and a list of the underlying assumptions finally is presented to the system fault tree analyst. The system analyst, the human reliability analyst, and someone involved directly with the performance of the activity of concern should



		EVENT	HEP	SOURCE*
A	=	CONTROL ROOM OPERATOR OMITS ORDERING THE FOLLOWING TASKS	.01 (.005 TO .05)	TABLE 20-22, ITEM 1 (P. 20-31)
B	=	OPERATOR OMITS VERIFYING THE POSITION OF MU-13	.01 (.005 TO .05)	TABLE 20-18, ITEM 3 (P. 20-28)
C	=	OPERATOR OMITS VERIFYING/ OPENING THE DH VALVES	.01 (.005 TO .05)	TABLE 20-18, ITEM 3 (P. 20-28)
D	=	OPERATOR OMITS ISOLATING THE DH PUMP ROOMS	.01 (.005 TO .05)	TABLE 20-18, ITEM 3 (P. 20-28)

* SWAIN AND GUTTMANN, 1983

FIGURE 6-8. ILLUSTRATIVE HRA EVENT TREE FOR ACTIONS PERFORMED OUTSIDE A CONTROL ROOM, WITH ESTIMATES OF NOMINAL HUMAN ERROR PROBABILITIES (NUCLEAR REGULATORY COMMISSION, 1983)



		EVENT	HEP*	SOURCE**
A	=	CONTROL ROOM OPERATOR OMITS ORDERING THE FOLLOWING TASKS	.02 (.01 TO .1)	TABLE 20-22, ITEM 1 (P. 20-31)
B	=	OPERATOR OMITS VERIFYING THE POSITION OF MU-13	.04 (.02 TO .2)	TABLE 20-18, ITEM 3 (P. 20-28)
C	=	OPERATOR OMITS VERIFYING/ OPENING THE DH VALVES	.04 (.02 TO .2)	TABLE 20-18, ITEM 3 (P. 20-28)
D	=	OPERATOR OMITS ISOLATING THE DH PUMP ROOMS	.04 (.02 TO .2)	TABLE 20-18, ITEM 3 (P. 20-28)

* THE HEP FOR EVENT A HAS BEEN MODIFIED TO REFLECT THE EFFECTS OF MODERATELY HIGH STRESS AND DEPENDENCE; THE HEPs FOR EVENTS B, C, AND D HAVE BEEN MODIFIED TO REFLECT THE EFFECTS OF MODERATELY HIGH STRESS AND PROTECTIVE CLOTHING.

** SWAIN AND GUTTMANN, 1983

FIGURE 6-9. ILLUSTRATIVE HRA EVENT TREE FOR ACTIONS PERFORMED OUTSIDE A CONTROL ROOM, WITH HUMAN ERROR PROBABILITIES MODIFIED TO REFLECT PSFs (NUCLEAR REGULATORY COMMISSION, 1983)

then review the HRA event tree and the associated assumptions. This is necessary to ensure that the human reliability analyst has correctly defined the success criteria for the system and that the system analyst does not apply the results of the HRA event tree outside the scope of its stated limitations.

6.6.3 OAT/TRC

THERP, as described in the previous section, is the standard technique for developing probabilities of errors in routine procedures ("slips") for fault tree applications. A leading methodology for the development of cognitive error ("mistake") probabilities is the Operator-Action Tree (OAT) with the associated Time Reliability Correlation (TRC) process (Hall, Fragola, and Wreathall, 1982). This methodology uses the time available to conduct a required action, versus a probability distribution of the time the action in fact takes, as the main basis for estimating the reliability of performance of the action. The primary concern is with estimating the probability of a operator being successful in diagnosing the need for, and ensuring the correct implementation of, necessary safety-related actions.

6.6.3.1 OAT/TRC Method

The OAT/TRC method involves three steps:

1. The development of the parameters of an OAT in terms of times available or required for the actions expressed in the tree, and the operator's delays in performing the actions.

2. The quantification of operator error probabilities by application of a TRC, with modifying factors where appropriate.

3. The transference of the quantified operator error probabilities to the system or other application fault trees (or event trees).

6.6.3.2 Structure of the OAT

The structure of an OAT is illustrated in Figure 6-10. Three potential failures that can result in a lack of timely and correct action in response to an undesired event:

- Failure to perceive that the event has occurred
- Failure to diagnose the event correctly and identify the necessary response to it
- Failure to implement the response correctly and in a timely manner.

Human reliability is then the overall probability of avoiding these three failures and arriving at a successful response to the undesired event. This probability is defined as the cumulative probability that a successful response is implemented by time T which is not greater than the maximum allowable time for the response, t. The component of this cumulative probability is the TRC which expresses the probability of not implementing a successful response by time t. The TRC parameters are estimated from observational and simulated psychometric data on operator actions. Note that the TRC function is analogous to the standard hardware time-to-failure distribution but in the human error context is assumed to be a log-normal rather than exponential or Weibull distribution. Then the parameters to be estimated are the median correct response time and the error factor (from which estimates for the log-normal's mean and standard deviation can be derived as described in Section 6.3.3.9). Dougherty and Fragola (1988) should be referred to for details on estimating the parameters from data and then adjusting the estimates in accordance with differences in operating conditions.

Uncertainties in the OAT/TRC modeling process (i.e., in the parameters of the TRC distribution) are integrated with the fundamental uncertainties in the time to respond, T, with an approximate multiplicative formulation. The random variable T is expressed as

$$T = t_R \times t_U \qquad \textbf{(6-76)}$$

where t_R is a log-normal variable with an estimated median and error factor as noted above; t_U is another log-normal variable with a median of 1.0 and an error factor that reflects uncertainty in the model for t_R. The error factor t_U's log-normal distribution is estimated judgmentally, supported by simulations and/or psychometric techniques.

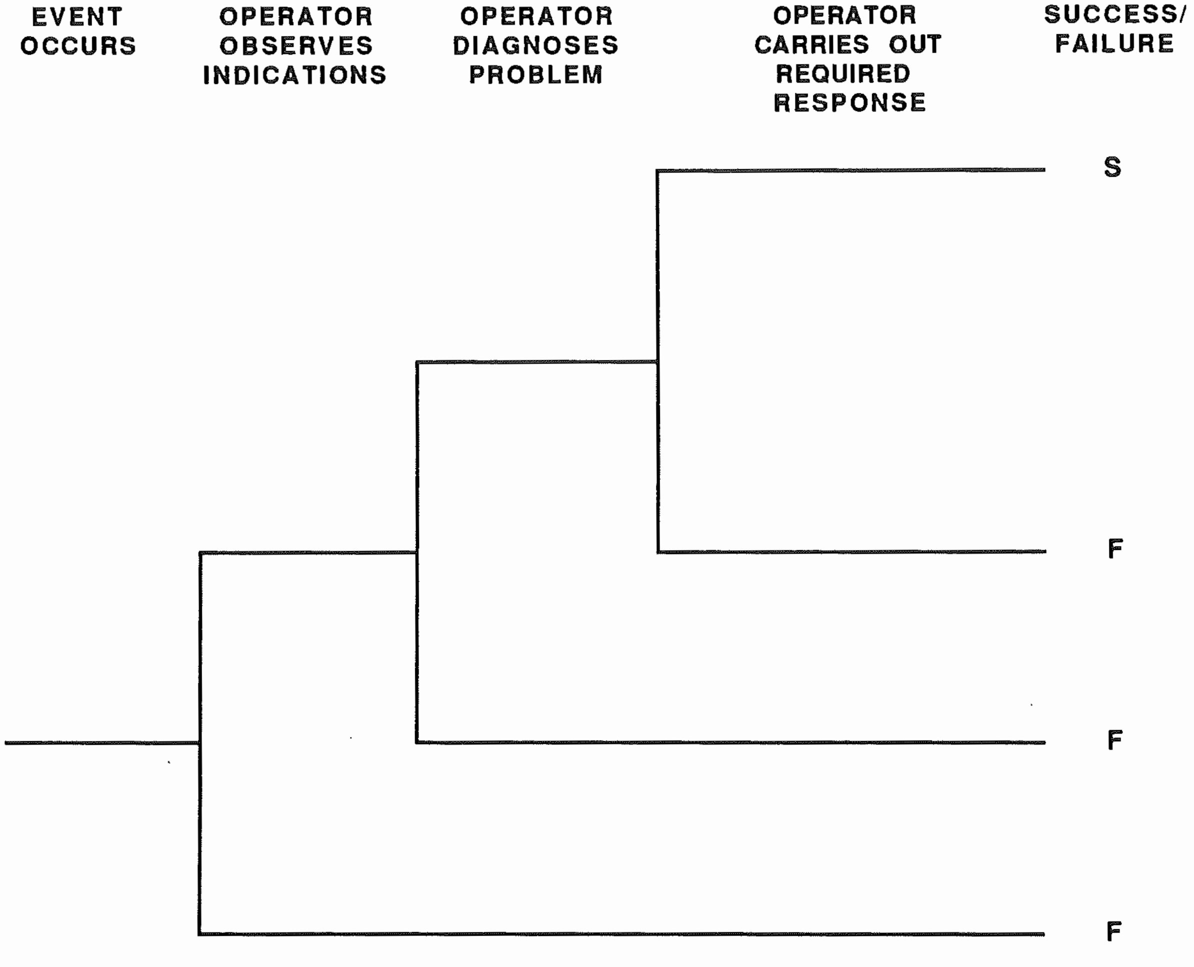

EVENT OCCURS
OPERATOR OBSERVES INDICATIONS
OPERATOR DIAGNOSES PROBLEM
OPERATOR CARRIES OUT REQUIRED RESPONSE
SUCCESS/ FAILURE
S
F
F
F

FIGURE 6-10. BASIC OPERATOR-ACTION TREE (WREATHALL, 1981)

6.6.4 Applications in NASA

The probabilities of cognitive errors or "mistakes" by operational or maintenance personnel in NASA's special activities are very likely to be negligible in relation to hardware failure probabilities. However, mechanistic errors or "slips," particularly by maintenance personnel, may have much more significant probabilities, especially during pressure-filled flight readiness activities. Based on analogies with personnel activities in the nuclear power industry, for which most HRAs have been carried out, probabilities of the order of 10^{-5} to 10^{-4} per mission may be reasonable for cognitive errors; 10^{-2} to 10^{-3} for mechanistic errors. THERP and OAT/TRC-type analyses need to be carried out, and experience gained in correlating their results with observations, to establish where HRA can and should contribute most usefully to fault tree models of NASA's hazardous activities.

CHAPTER 7
RISK EVALUATION FOR ACCEPTABILITY OR MITIGATION DECISIONS

This section presents concepts and generic criteria for supporting judgments on whether a predicted risk level for a hazardous activity in NASA, as well as generally, is sufficiently low for the activity to be instituted or continued, or whether mitigation measures may be required, feasible, and beneficial.

These considerations relate to such questions as:

- How safe is a particular hazardous activity?

- How does this safety compare with the safety of other activities?

- How much additional safety could be attained for a given cost, through some set of alternative modifications?

- How much would it cost to attain some required level of safety, through some set of alternative modifications?

- Which would be the safest means of achieving a given objective (e.g., transport of a given amount of a given material in a year over alternative routes or by alternative modes or by alternative shipment sizes)?

- How much added risk would be imposed in some other activity due to a modification or alternative that decreases the risk in a given activity (e.g., energy from coal instead of nuclear power will cause more rail crossing accidents, more coal miner deaths and illnesses)?

- Central socio-political, or programmatic, issue: Is the estimated (perceived?) risk "acceptable"? What are ways of appraising this?

7.1 RISK ACCEPTABILITY EVALUATION

While no single approach has yet been established that enables a universally appreciated evaluation of the acceptability of the risk of a hazardous activity (see Lowrance, 1976, for a review of the concepts involved), a number of attempts have been made to develop such an approach. These are discussed here in four categories: De minimis and ALARA criteria, comparisons to "ambient" or historically accepted risks, comparisons with risks of equi-benefit alternatives, and balancing of risks and benefits.

7.1.1 De Minimis and ALARA Criteria

These two kinds of criteria for acceptable risks apply when it is judged by whatever means are available that it is not worthwhile to attempt to reduce the subject risks and thus, ipso facto, they are acceptable. De minimis criteria (see Fiksel, 1985) establish this by considering that the risks are negligible; no significant likelihood of significant harm exists. ALARA (As Low As Reasonably Achievable) criteria (see Higson, 1985) apply when it is judged that all economically and technologically practicable efforts have been made to reduce the risks to their existing levels. (In the United Kingdom, closely related ALARP (As Low As Reasonably Practicable) criteria are employed, differing primarily only in their application to risks to individuals rather than population groups.)

It seems clear that what is de minimis or ALARA to one person may not be to another. The application in practice of these criteria requires sufficient authority to accept their implications.

7.1.2 Comparison to Ambient or Historical Risks

In 1969, Chauncey Starr published the first of many articles on public risk acceptance in relation to benefits, as revealed by historical data (Starr, 1969). Expected fatalities per hour or per year and per individual in various groups exposed, due to voluntary or involuntary hazardous activities, to potential accidents and other deleterious elements were estimated from past data and then compared to assessments of the benefits accruing from these activities (see Figure 7-1). Starr found that historical levels of risk acceptance increased roughly in proportion to the cube of the increase in benefits, and that voluntary acceptance levels were about three orders of magnitude greater than involuntary acceptance levels. (These particular conclusions have since been disputed, however (Otway and Cohen, 1975).)

Starr's concepts have been extended by many others in attempts to establish numerical acceptable risk levels for hazardous activities such as petrochemical and energy facilities that provide specific benefits or meet specified societal needs. These numerical levels may also reflect the confidence in the risk estimates that are evaluated (Okrent and Whipple, 1977).

Three major philosophical problems exist with the approach to risk acceptability evaluation based on Starr's concepts. First, for involuntary risks, the groups accepting the risks often differ from the groups receiving the benefits (or at least do not share the benefits in a manner reflecting their exposure to the added risks). Second, the use of a risk measure based on expected, or average, or mean, losses, while convenient, forgoes any ability to distinguish low probability/high consequence from higher probability/lower consequence risks. The former are often of more critical concern to the public and other decision-makers. The societal "disutility" of accidents appears clearly to be nonlinear as accident magnitude increases. The utility functions to express this have been discussed, but they have not yet been developed meaningfully. Finally, the groups evaluating the risks of a hazardous activity may differ greatly in their perceptions of its benefits as well as its risks, and thus differ on the acceptability of the activity.

Several psychometric experiments have been reported which attempt to assess how individuals balance their perceptions of the risks and benefits of hazardous activities. While consistent with Starr's generic results in some aspects, great differences were also exhibited, depending on the availability to individuals of information on the activities, their familiarity (or their beliefs that they were familiar) with these activities, and so on. The problem of obtaining a consensus on the acceptance of risks to provide specified benefits is evidently one that is very difficult to resolve (Slovic and Fischoff, 1979).

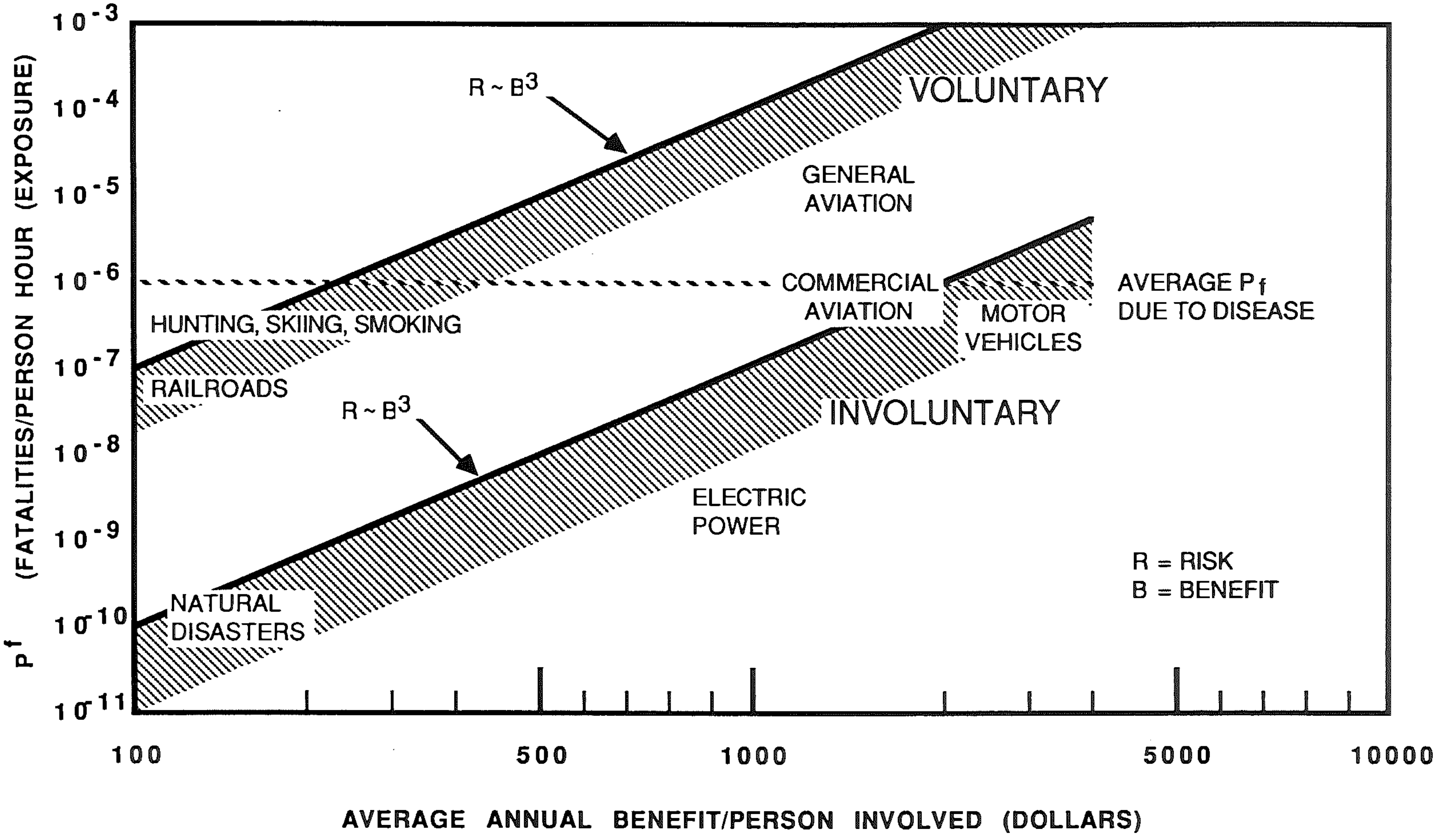


FIGURE 7-1. ACCEPTED RISK VERSUS BENEFIT, VOLUNTARY AND INVOLUNTARY EXPOSURE (STARR, 1969)

The second of the philosophical problems noted above is the only one that so far has been meaningfully attacked. An example is provided by the well-known attempt at risk acceptability evaluation (albeit not presented in such terms explicitly) in the Nuclear Regulatory Commission's Reactor Safety Study (1975). Complete risk profiles reflecting the probability distributions of all possible losses, rather than only their means, are generated for nuclear power plants and compared to the profiles for various ambient and historical hazards, natural and man-made.

The principal weaknesses of the ambient/historical risks comparison method (over and above arguments on the validity of the profile functions developed) is its neglect of the fact that, even if the incremental risk of the hazardous activity is small compared to the total ambient risk, the proposed involuntary risk-takers do not always happily accede to even the small addition. Overcoming this attitude, when it is justified to do so, can be a major problem. All risk evaluation procedures imply that this can best be done by increasing the risk-takers' benefits (real or perceived). Secondarily, any means for enhancing the credibility of the risk estimates to them would be helpful, but probably not decisive.

7.1.3 Risk Comparisons of Equi-Benefit Alternatives

A second risk acceptability evaluation approach is the standard operations research technique of assuming that some activity must be put in place to satisfy a specific need, and then establishing which alternative means of implementing it would give rise to the least risk. On this basis, for example, nuclear power has been argued to be safer overall than coal for generating electricity (taking into account only the mean values of the two risk profiles and employing, to some extent controversial, "accounting" of total system risks from raw material mining to energy production).

On the surface, the procedure should be a strong one for not merely evaluating, but also encouraging, reasonable risk acceptance. However, sometimes no practical alternative is deemed acceptable to the public or its spokesmen. They may demand some approach based on unproven or uneconomic technology, or the avoidance of the needed activity entirely (even at some unconsidered other risks). Nevertheless, this method, perhaps combined with procedures for determining the incremental benefits necessary to induce rational risk acceptance, may be the most suitable for some hazardous activities of importance.

7.1.4 Balancing of Risks and Benefits

Quantitative procedures exist for expressing the risks of a hazardous activity, as well as its benefits, in common economic terms, e.g., present-value dollars. However, these procedures generally entail assuming or inputting a "value-of-a-life," and it has been difficult to obtain agreement on this feature of the analysis (see, e.g., Linnerooth, 1975). If an agreement were possible, it could then be argued that a hazardous activity was acceptable if the potential expected loss induced by its risks were less than the dollar value (or some fraction of this value) of its potential benefits.

7.2 EVALUATION OF CANDIDATE RISK MITIGATION MEASURES

Mitigation measures may reduce a risk by reducing the probability of occurrence of a mishap, or by reducing its potential losses if it should occur. Potential loss reductions may result from diminishing the probabilities of the higher intensity effects (e.g., by using materials with less explosive potential), by minimizing potential consequences (e.g., by strengthening structures protecting people, or by providing sheltering or escape facilities), or by reducing maximum loss potential (e.g., by evacuating people, or by operating only in areas of low population density).

7.2.1 Mitigation Effectiveness Prediction

It is often difficult to assess the effectiveness of candidate mitigation measures, to evaluate them comparatively, and to select one that is most cost-effective. This difficulty arises especially when the estimates of the risks of concern have been developed from statistical data on past mishaps and loss experience. The effectiveness of a mitigation measure must then be predicted as the result of the hypothetical effect on the past data that would have accrued if the mitigation measure had been in place while the data were being acquired. Evidently, judgment is an essential factor in such a prediction.

As was noted earlier in the discussion of logic trees, however, if a fault tree in sufficient detail can be successfully applied to the mishap analysis, a more straightforward procedure becomes available for predicting the decrease in the risk resulting from at least those mitigation measures that focus on the mishap's occurrence probability. It is then only necessary to recalculate the probability of the mishap, given that the particular mitigation measure has been applied to the elements of some of the fault tree's event sequences describing the possible mishap occurrence modes, thereby eliminating or decreasing the probabilities of such modes.

7.2.2 Cost-Effectiveness and Cost-Benefit Evaluations of Mitigations

As has been indicated, the effectiveness of a risk mitigation measure is quantified by the reduction in risk (the "Δ risk") it provides. This reduction may be assessed in terms of an expected loss averted, or in terms of more comprehensive differences between the relevant risk profiles, with and without the mitigation measure. The effectiveness of alternative measures that can be implemented within available financial and other resources can then be compared, and the alternative selected that provides the greatest effectiveness. Similarly, an alternative could be selected from all those considered to meet a given risk reduction requirement as the mitigation measure of lowest cost.

A related approach is the comparison of the cost of a risk mitigation measure with the decrease in an expected loss or other desirable change in a loss probability distribution. For example, if the loss in longevity in the population near a hazardous activity is of concern, the effectiveness of a mitigation would be the increase in longevity that it would induce in the population exposed to the risk. Schwing (1979) constructs an effectiveness index defined by the cost of a particular life-extending program divided by the longevity increase it provides. The index is then the cost in dollars to gain a year of longevity for the population affected. A scheme such as this for the evaluation of the cost-effectiveness of alternatives has the advantage that it not only places the costs of various mitigation measures in relationship to one another, but enables these costs to be put in perspective with safety expenditures in other sectors of an agency, an industry, or society as a whole.

A complete implementation of a cost-effectiveness approach requires a realistic accounting of all costs (and other "dis-benefits"). Besides the direct costs of an alternative, which include capital, operation, and maintenance costs, the costs of mission delays and degradations and other indirect costs may also need to be incorporated. A utility theory approach to such a multi-attribute decision problem may be applicable (see, e.g., Keeney, 1980).

In addition to the estimation of the effectiveness of risk mitigations, as in the foregoing procedure, it is sometimes also important to estimate their benefits; that is, the translation into economic terms of the value of the reductions in risk they provide. This is required for many areas of Federal Government safety regulations by Executive Order 12991 (1981), for example. The purpose is to justify a mitigation by exhibiting that its cost is exceeded by its benefits expressed in common terms. When human safety is involved, this can give rise to the need to establish a value for a life saved, often a controversial consideration, or, less controversially, the marginal value of the mitigation-caused decrease in the chance of a loss of life due to this hazardous activity.

7.2.3 Residual Risks Control

After the risk management actions that so far have been discussed are applied to a hazard, some level of residual risk will remain. This level may be deemed acceptable in accordance with such criteria as have been described, with no further action required. Alternatively, the residual risk may have to be tolerated for sufficient technological, economic, or mission importance reasons. And, it may nevertheless be desirable to take whatever additional actions as are possible to ameliorate the impacts of an associated mishap should it occur.

When a higher risk than one that is fully acceptable is tolerated in NASA (as also in the military, and usually when it is considered to be at least a low-probability risk), a waiver or variance is established by a management level with the legal authority to do so. It may be noted that it has not always been evident to the waiving authority exactly what the risk was that was being accepted through the waiver; certainly the risk's quantitative level has rarely been known. An important function of risk management based on quantitative procedures is to provide to the waiver decision-maker an assessment of the level of resulting accepted risk which is as accurate as possible, including consideration of the uncertainties that are present in the assessment.

Given the acceptance or tolerance of a residual risk, it may yet be possible to establish contingency plans and plans for emergency actions that can be expected to diminish the harm from the mishap or mishaps reflected in the risk. This, in fact, represents a means for reducing the risk by reducing the likely severity of the mishap's consequences, but it is not usually the case that this reduction can be quantified in a meaningful way, in part because of the variability and uncertainty in the execution of a contingency or emergency action that can be expected. Emergency plans, such as are developed for fighting facility fires, for example, may require the participation of external emergency forces, as well, and this adds to the uncertainty on the efficacy of the emergency actions.

Intimately related to the contingency plans and plans for emergency actions is the definition and implementation of emergency systems and procedures, and the training of the personnel who will employ them. Included in these systems and procedures are capabilities for detecting precursors of mishaps and other problems, as feasible, and for providing warning to operational and emergency personnel. They include escape systems, when feasible, sheltering capabilities, and evacuation procedures and resources. Risk management is concerned with the definition and evaluation of alternatives for these emergency systems and procedures and the assurance of their proper implementation.

7.3 CONSTRAINTS ON RISK MANAGEMENT APPLICATIONS

In the discussions in this document of risk estimation and risk evaluation procedures, various problems constraining their application in risk management decision-making have been recognized. Primarily, these problems have to do with, first, the uncertainties arising in the risk estimation procedures, and, second, the policy issues arising in the use of the estimates in safety decision-making. Secondarily, the availability of the necessary technical and financial resources to support the risk management process may also be a problem.

7.3.1 Resource Requirements

The latter problem, the availability of technical and financial resources, must of course be addressed in the delineation of each Safety Risk Management Program Plan. Maximum effect in NASA's applications of risk management would accrue from its uniform employment in all significant safety decisions irrespective of cost considerations. Lower resource requirements would have to be met if risk management in NASA were limited to particular kinds of decisions whose supporting risk analyses were conducted only by a special organization. A program plan should reflect the approach that is best in the spectrum of possibilities bounded by these two extremes.

7.3.2 Uncertainties

Returning now to the primary problems, consider first that of uncertainties. This problem underlies the lack of full acceptance of quantitative risk analysis as a basic tool in safety decision-making (Philipson and Gasca, 1982). The essential source of the uncertainties is in the shortcomings (both inherent and correctible with added effort) of the data available for risk estimates, most particularly for the low probability mishaps of most significance (Parry and Winter, 1981). Many attempts have been made to augment the data bases for such estimates. As has been noted earlier, these have ranged from the crude use of surrogate data from "similar" contexts in transportation risk analyses (see, e.g., Fairley, 1975) to a much more sophisticated and statistically correct use of such data in nuclear power plant risk analyses (Martz and Bryson, 1982).

The most practical means for overcoming data problems, however, is to change the data requirements through modeling. Instead of system- or activity-level mishap data, for which for most important hazards the record is inherently sparse for the rare mishaps of greatest interest, combinatorial models require data only on the occurrence of individual events leading to the mishap events that may be common to a wider class of mishap occurrences, and perhaps also to tests in which the final, harmful event in the particular mishap sequence of interest does not occur. Even more generally, fault tree models transform the system-level data requirements to requirements for data on failures and associated opportunities for failures (actions, exposure times or distances, etc.) associated with equipment and human components. Such component data can be readily available. Although it is not known to have been exploited significantly as yet, fault trees also lend themselves to the possible employment of "close-call" data, that may support quantification of the probabilities of events that contribute to a mishap (Philipson and Gasca, 1982).

Given the best that can be done to establish a good data base and implement effective modeling, the risk estimates ultimately attained will nevertheless usually reflect important uncertainties. As was seen in Section 5, above, various methods are available to quantify at least the uncertainties in the risk estimates that arise from known or (relatively straightforwardly) assumed uncertainties in the data (Martz et al., 1983). Uncertainties in the risk estimates that arise from uncertainties in the modeling assumptions are much less readily handled. Possibly effective, if costly, means of doing so would be independent repetitions of the risk analyses, and comparisons of the methods by which they arrive at any differences in their results (see, for instance, Philipson, 1982). Of course, a rigorous dissection of an analysis by impartial experts could lead to a similar end--a reasonable degree of confidence in the modeling results.

The "bottom line" in dealing with uncertainties is that their magnitudes, when possible, and their known and potential significances, always, must be clearly expressed to the safety decision-maker(s) for whom the risk analysis was performed. The effective communication of risk analysis results is thus as important a function of risk management as the development of the results in the first place. When legal issues are involved, such effective communication may also be required to less technically aware members of the courts and of legislatures, as well. And, of course, effective communication with the public media may be essential to the continuance of the hazardous activity of concern.

A more comprehensive discussion of the impacts and treatments of uncertainties in risk decision-making is provided in Section 8, following.

CHAPTER 8
THE IMPACTS AND TREATMENT OF UNCERTAINTIES IN RISK DECISION-MAKING*

In this chapter will be considered how uncertainties in the risk estimation function of risk assessment affect the risk evaluation function and what can be, or potentially could be, done to adapt to these uncertainties more effectively in risk management decision-making employing risk assessments. This will be carried out for the several kinds of risk evaluation techniques in turn:

- Comparisons to ambient risks (including the establishment of generic numerical acceptance criteria)

- Comparisons to revealed or expressed preferences

- Risk-cost-benefit evaluations.

It is intended that the present appraisal of the sources, characteristics, and impacts of uncertainties in risk assessment will motivate and provide a basis for deeper investigations of specific problems in the future. In particular, the specific quantifications of uncertainty discussed here are recognized to sometimes exceed present capabilities. Nevertheless, they provide, as a minimum, an ideal against which present capabilities can be measured, and towards which further uncertainty analysis efforts should be directed.

8.1 UNCERTAINTIES IN RISK ASSESSMENT: GENERAL

As has been indicated, uncertainties in any complex modeling procedure, such as the alternative procedures usable in risk estimation, arise from incompleteness and inaccuracy in the data available to the modeling, inexactitude imposed by the assumptions and techniques employed to overcome the data's shortcomings as well as fundamental inadequacies in the understanding of the physical processes being treated, and finally, of course, the elements of real randomness in these processes. The last source of uncertainty immediately imposes

* This chapter draws primarily on material presented in Philipson (1982).

the use of probabilistic models, central in all risk analyses. The choices of the distributions, parameters, and parameter values employed in these models also respond to models or views of the other uncertainty sources as well, and the means available for their investigation and, as feasible, resolution.

It is of value to appreciate in this that what is being accomplished in a risk estimation is the translation of objective sample data and/or subjectively-derived information into an inference for a "true" or, in statistical terms, "population" probability function (that may be specified by one or more population parameters). The understanding of the uncertainties in this procedure (beyond the fundamental randomness) translates into the relationship established between the derived estimate and the true function or parameters.

Roughly speaking, uncertainties "add" to the risk estimates produced by a nominal model for the true risk. Thus, for example, if a risk for some hazardous activity is estimated as an expected loss (in each loss dimension of interest, such as fatalities, injuries,...), the uncertainties in the probabilities and losses that integrate into this estimated expected loss imply an uncertainty range of the corresponding true expected loss. The magnitude of this range might be given deterministically as an expected loss increment over the estimated value that is due to the worst-case combination of uncertainties in the contributing probabilities and losses, or it can itself be probabilistic, e.g., in terms of a confidence interval, if the contributing uncertainties are given probabilistically. Risk evaluations employing expected losses as the measure of risk, e.g., in terms of expected fatalities per exposed individual per year in comparisons to ambient expected fatalities as in Figure 8-1, could reasonably then use an upper confidence limit on the expected loss (and, where uncertainties are also present in the ambient risk estimates, lower confidence limits for them) as a conservative value in the comparisons. When formal statistical procedures cannot be used to develop such confidence limits, qualitative and subjective appraisals of the uncertainties in the estimation of the expected loss are nevertheless often able to provide rough "high-confidence" upper limits. Of course, this may excessively increase the conservativeness of the results if care is not exercised to maintain "reasonableness" and "credibility" throughout.

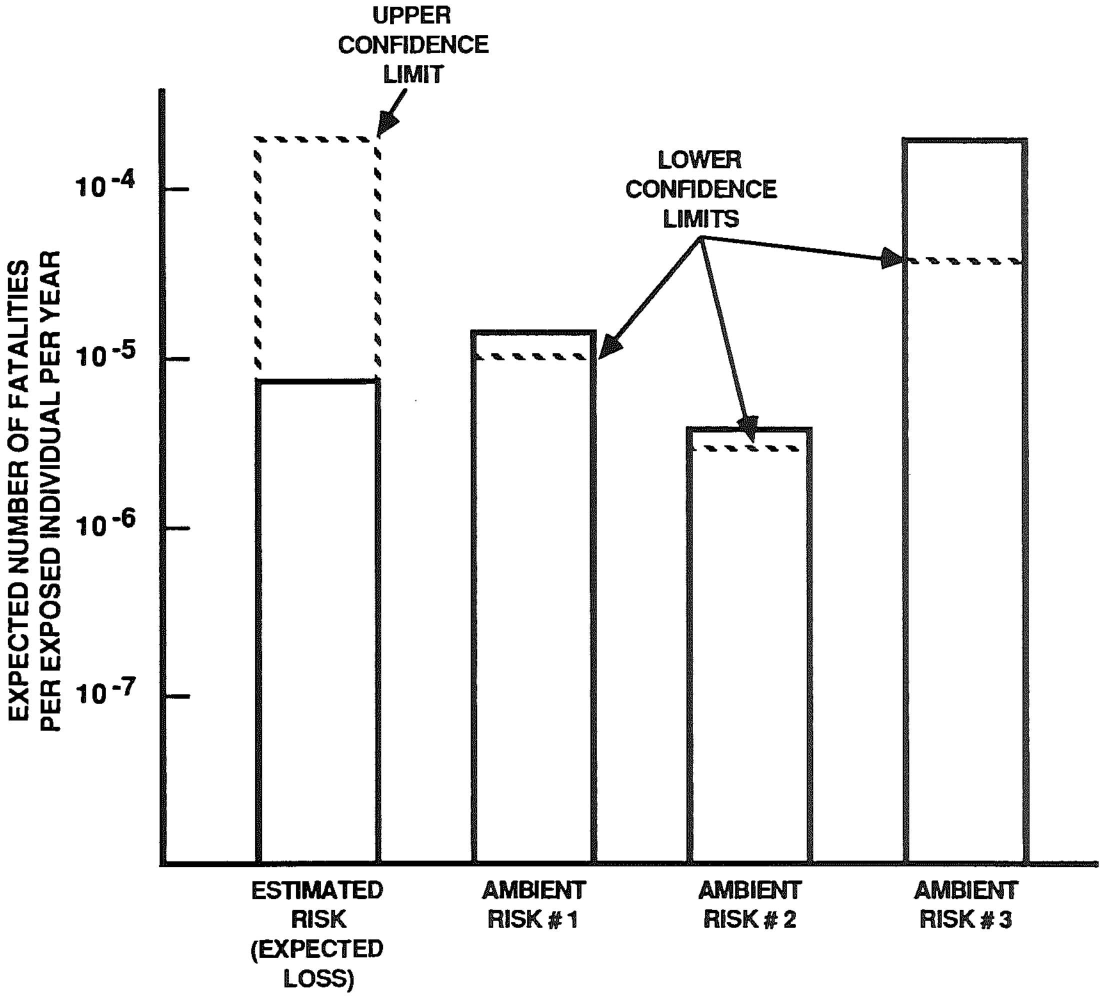


FIGURE 8-1. IMPACT OF UNCERTAINTIES IN COMPARISONS OF ESTIMATED TO AMBIENT RISKS (EXPECTED LOSSES)

If a complete risk profile rather than only an expected loss value is modeled, the contributing uncertainties now affect, in more or less complex ways, the individual parameters of the distribution function corresponding to the profile.* The result is a probabilistic range of variation in each of these parameters and a corresponding range of variation for any value of the profile. Thus, the probability, nominally 10^{-4}, say, of exceeding 10 fatalities in the life of the hazardous activity that gives rise to the profile, may now become stated as a range of values such as 10^{-5} to 10^{-3}, with some given confidence, say 90%. The conservative upper confidence limit, 10^{-3}, might be used alone and, as this is done for various possible loss values, an upper "bounding profile," expressing an upper 90% confidence limit for the probability of exceeding each possible loss value, would be generated.** A selected set of resulting risk values such as the probability of 10^{-3} of exceeding 10 fatalities might then be compared one by one to the corresponding probabilities from relevant ambient hazards in an evaluation of the significance of the risk profile. This is illustrated in Figure 8-2a.

Another way of introducing the effects of uncertainties into a risk profile (that retains its probability properties) is that of Bayesian revision of the nominal distribution function associated with the profile. Figure 8-2b illustrates this approach. The nominal distribution could derive from limited objective data, e.g., a sample mean and variance defining an assumed normal distribution for the potential loss values. In the Bayesian procedure, the uncertainty in, say, the true mean of the distribution could be given itself as a specific subjectively-defined normal distribution for the possible values of the mean. Combining the nominal and uncertainty distributions in this case would lead to a revised distribution that would also be normal and reflect both the objective and subjective information available. This distribution's associated risk profile would then be used as before; e.g., for comparison to ambient risks of various magnitude losses, such as in the linear ambient risk profile shown in Figure 8-2b.

* Recall that the risk profile for a given loss dimension is the complement of the cumulative distribution function: $P(X \geq x) = 1-P(X \leq x)$ where x is any given value of loss.

** It should be noted, however, that this bounding "profile" generally will no longer be the complement of a cumulative probability distribution function. To retain this property, a procedure like that next described would be necessary.

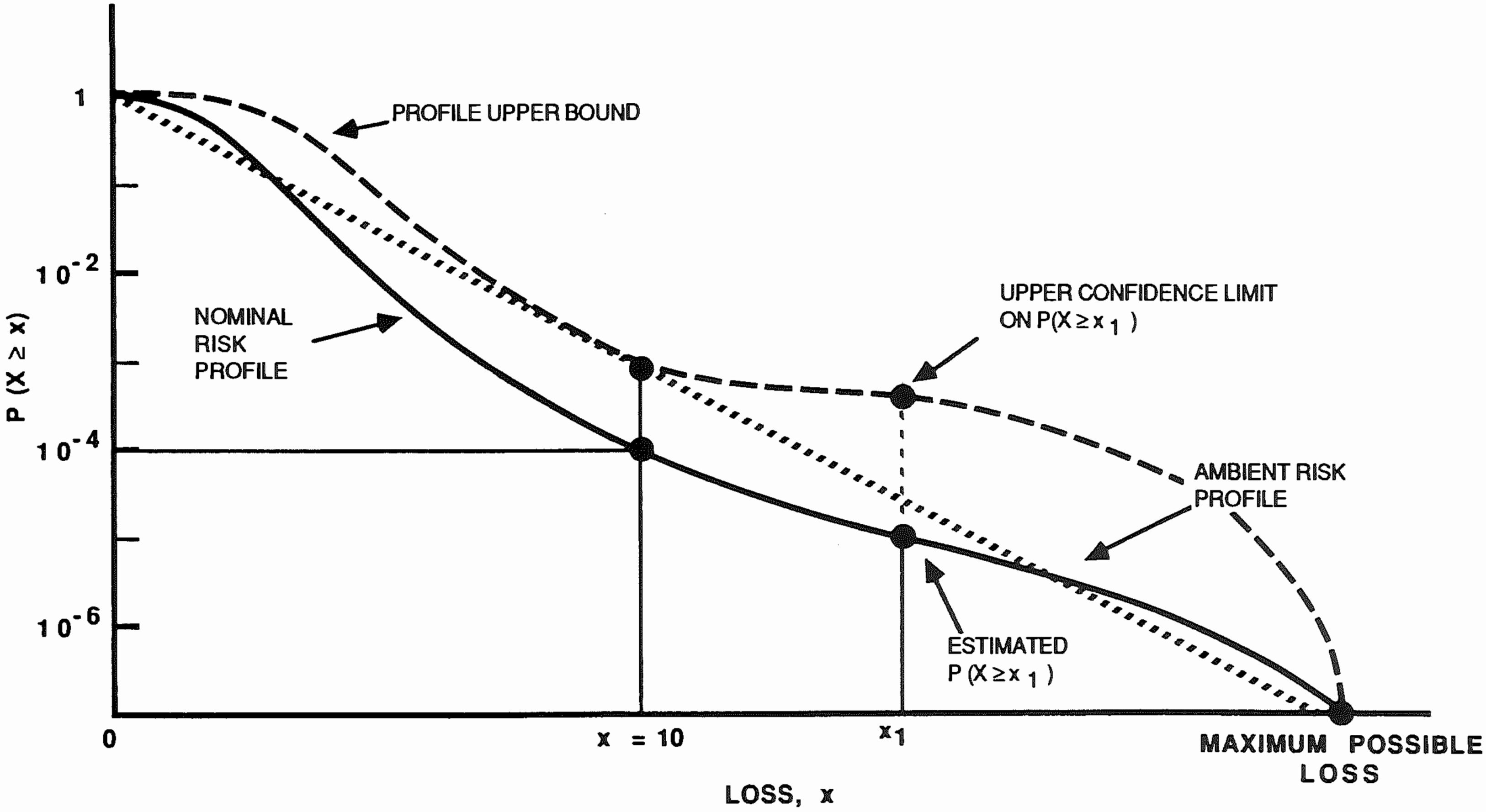


FIGURE 8-2a. IMPACT OF UNCERTAINTY ON RISK PROFILE (FIRST APPROACH)

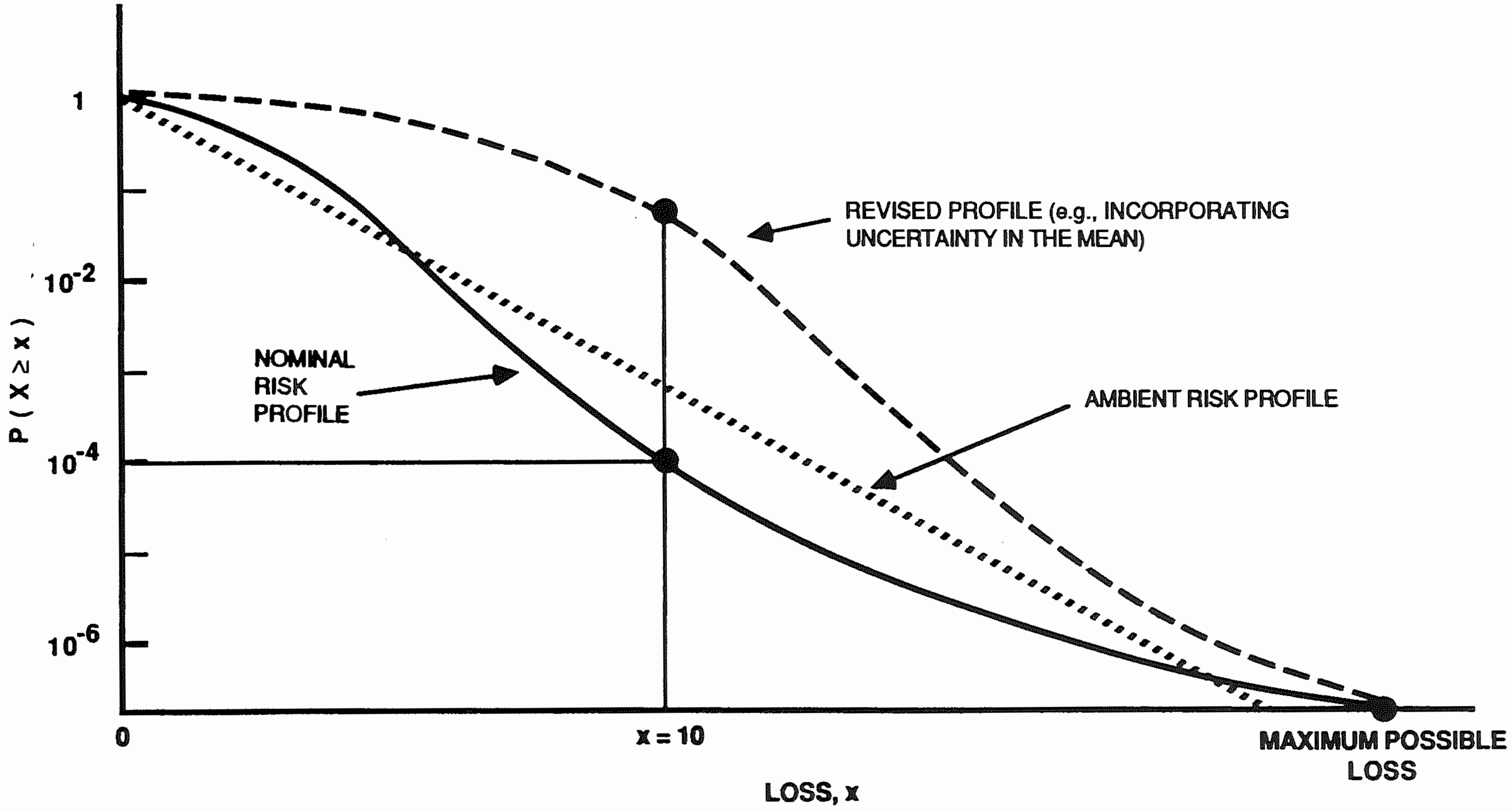


FIGURE 8-2b. IMPACT OF UNCERTAINTY ON RISK PROFILE (BAYESIAN APPROACH)

The most general approach to introducing probabilistic uncertainties into a risk profile is to replace its associated probability distribution function with the joint distribution of the loss variable(s) and variable(s) expressing the uncertainties, and obtain thereby a joint risk profile. The Bayesian approach accomplishes this in a specialized form, where the uncertainties are specific to known parameters of the loss variable's distribution. The general procedure is not yet believed to be of more than theoretical interest. If it were carried out, however, it would produce an "upper bound" risk profile similar to that shown in Figure 8-2b.

Finally, it is to be recognized that uncertainties cannot always meaningfully be made explicit, even by subjective appraisals. In such cases, risk evaluations can only introduce the impacts of such uncertainties as greater or lesser degrees of required stringency in the interpretation of the significance of the estimated risks. Thus, for example, if an estimated risk is several orders of magnitude smaller than all ambient risks, and if it cannot be seen how it could approach them in any reasonable view of the accuracy of its estimation process, then the estimated risk could be deemed acceptable (with an ambient risks comparison criterion). If, on the other hand, the estimate is not much smaller than some ambient risk, the view could be taken that it might reasonably, in fact, exceed the ambient risk and so could not be decided to be acceptable.

The foregoing general concepts are next applied with some specificity to the three kinds of risk evaluation techniques.

8.2 UNCERTAINTIES IN RISK EVALUATIONS BASED ON COMPARISONS TO AMBIENT RISKS

As the examples given in the preceding general discussion show, the introduction of considerations of uncertainty in risk evaluations that are based on comparisons of an estimated risk to ambient risks is relatively straightforward. A set of individual ambient risks given as expected losses from various extant hazards provides a basis against which a new risk can be compared. If the new risk is not significantly less (after its upper bound, and, if necessary, the ambient risks' lower bounds are taken into account), then it is deemed unacceptable.

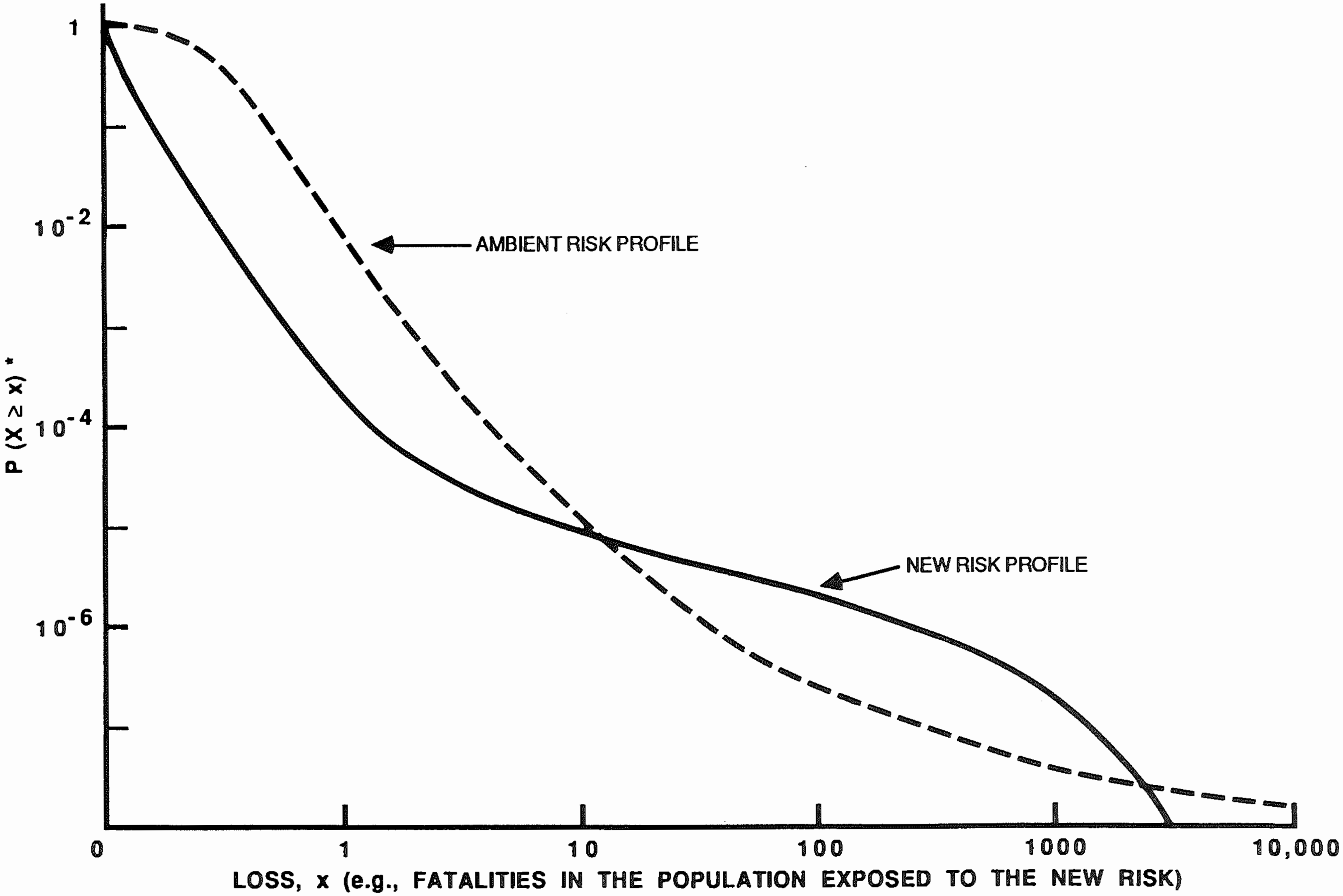
A qualitatively equivalent procedure is to assign a set of uncertainty factors (less than unity) to a nominal ambient risk level expressed as an expected loss. Here the factors reflect the possible degrees of uncertainty in the estimation of a new risk. Then it is required for its acceptability that the estimated new risk be less than the ambient level times the factor appropriate to the uncertainty in the estimate. Thus, for example, if a new risk's estimate were highly uncertain, it might be required to be less than a criterion value of one-tenth of the ambient risk level. If it were only moderately uncertain, its being less than one-half of the ambient level might suffice for its acceptability.

An analogous procedure allows an amplification in a nominal acceptable risk level that reflects the qualitative benefits in acceptance of a new risk. The greater the benefits, the larger could be the amplification factor.

Finally, if the new risk is of a catastrophic nature, its expected loss value might be amplified by raising the possible loss magnitude to some power such as two or three, prior to calculating the expected value. This amplification reflects the non-linear disutility to society of larger accidents. This, of course, results in such accidents' risks becoming less able to compare favorably to the ambient risks than if the same expected loss value were to accumulate from a greater number of smaller potential accidents.

All of the foregoing risk acceptability criteria definitions would operate with risk expressed as a single numeric: the expected value of loss, possibly biased to reflect various qualitative assessments of the precision of the risk estimate, the importance of the benefits from its acceptance, and the extent to which it derives from potential catastrophes. When, however, the distinct risks of different magnitude losses, rather than a single expected loss, are considered, the comparison to ambient risks should be made separately for the different magnitudes. Or, if the data permit, a complete risk profile could be taken into account and compared to a set of ambient risks of losses of different magnitudes or, possibly, to an integrated ambient risk profile (such as the hypothetical linear ones in Figures 8-2a and b). Modifications of the new risk's or the ambient risk's profiles to reflect uncertainties can be taken into account in this.

It remains to consider the decision procedure this multi-attribute (discrete or continuous over the different possible loss values) comparison entails. Consider, for example, profiles related as in Figure 8-3. (It is assumed that the profiles are already revised, as necessary, to reflect the levels of confidence in their estimation.) It is seen that the ambient risk profile



* Probability of loss exceeding a value x, per year (say), for a new hazardous activity, or for accumulated ambient risks. The profiles are assumed to be appropriate bounds reflecting estimation uncertainties.

FIGURE 8-3. COMPARABLE NEW AND AMBIENT RISK PROFILES

exhibits higher probabilities of occurrences of moderate losses (up to x=10), and also of very great losses (more than, say, 5000 due, e.g., to the remote possibility of very great natural catastrophes such as an extreme earthquake or the fall of a meteorite on a city) beyond the possible range for the new hazardous activity. The new risk's profile, on the other hand, has significantly higher probabilities of losses exceeding values in the intermediate-to-large range of, say, 10 to a few thousand. Note that it is quite possible that the new expected loss could be much less than the ambient expected loss. Should the new risk be accepted, on the basis of comparison with the ambient risk?

Clearly, the new risk's acceptability depends on how the decision maker(s) assesses the relative importance of the different loss ranges. He may make this assessment and apply it to the two curves entirely judgmentally, and decide, for instance, that the admittedly more likely low losses and the very unlikely extreme losses in the ambient profile are not of as much concern as the moderately likely intermediate size losses in the new risk profile, and so decide against the acceptability of the latter.

A more quantitative approach could be adopted. Relative weights could be assigned for the different ranges between loss values where dominance switches between the two profiles. The mean value of loss for each profile in each range could then be calculated and the weighted sum of these values for each profile then compared. For example, suppose (unrelated to Figure 8-3) the mean values and the subjective importance weights of the several ranges are as follows:

		AMBIENT RISK		NEW RISK	
LOSS RANGE	WEIGHT	MEAN	MEAN x WEIGHT	MEAN	MEAN x WEIGHT
0-10	1	10^{-2}	10^{-2}	10^{-4}	10^{-4}
11-5,000	100	10^{-6}	10^{-4}	10^{-5}	10^{-3}
>5,000	10,000	10^{-6}	10^{-2}	0	0
TOTAL			$2x10^{-2}$		$1.1x10^{-3}$

The decision-maker might now decide in favor of the acceptability of the new risk.

It is clear, however, that this procedure would be better conducted if the several ranges were further broken down into more, smaller ones. In the limit, what would be used would be the expected value for each profile of the weighted loss, with the weight given as a continuous function. The expected values would then be obtained by an integration of the weighted loss times the probability density function (given by the negative derivative of the profile curve) for the new and ambient hazards' potential losses. The result would be a comparison of the decision-maker's "expected disutilities" for the two risks, with his disutility function defined by the subjective weighting function. The mathematical formulation of this procedure is:

New risk profile p_N $(X \leq x)$ is acceptable if

$$\int_{o}^{\infty} w(x) \bullet p_N (x)\, dx < \int_{o}^{\infty} w(x) \bullet p_A (x)\, dx \qquad \textbf{(8-1)}$$

where

$w(x)$ = Decision-maker's weighting function for the relative importance of a loss of magnitude x (in the exposed population and during a year, say, of exposure)

$p_N(x)$ = Probability density function for the new risk

$p_A(x)$ = Probability density function for the ambient risk

This procedure is also directly extendable to the consideration of several different types of losses such as injuries of different severities as well as fatalities, delayed fatalities (e.g., due to cancers resulting from exposures to carcinogens) for different periods of delay, health degradations, property losses, environmental damage, fatalities and injuries to different exposed groups (the public, workers, ...), etc. It can also be further formalized with more concrete and less arbitrary means of establishment of the disutility function. On the other hand, the procedure outlined can also be specialized to any simpler form of acceptability decision-making on the basis of comparison of any characteristics of the new and ambient risk profiles, where these have been revised as necessary to reflect the uncertainties in the new and ambient risk estimates.

8.3 UNCERTAINTIES IN RISK EVALUATIONS BASED ON COMPARISONS TO REVEALED OR EXPRESSED RISK PREFERENCES

Instead of risk profiles, the risk, such as the estimated expected loss (for any given loss dimension), and the benefit (in expected dollars or other measure of value) accruing to an exposed group from a new hazardous activity, are now compared to the corresponding risk and benefit of classes of other activities that have been accepted in the past or are now otherwise expressed to be acceptable by a decision-maker.

The past or expressed acceptance of the latter activities is assumed to reflect the past or present risk-benefit preferences of the exposed group.

Starr's two accepted risk versus benefit curves (recall Figure 7-1) are the original representations of these preferences, for voluntary and involuntary risk-taking groups, respectively. Analogous curves could in principle be established from assessments of expressed risk-benefit preferences. If a new voluntary or involuntary risk's expected loss versus benefit point (established objectively from data, or subjectively from attitudinal assessments) falls below a corresponding preference curve, it is presumptively acceptable on the basis of the present evaluation procedure.

If several loss dimensions are of concern, the new risk would be deemed acceptable only if all of its component expected loss-versus-benefit points fall below individual accepted expected loss-versus-benefit curves established from the past record for each loss component. (Such curves are not now known to exist except for the fatalities component, but curves for injuries and property damage, at least, appear able to be developed if desired.)

Note also that instead of expected loss, curves could in principle be established for some exceedance probability from a risk profile, $P(X \geq x)$, for a given loss x to provide the basis for comparison of the new risk's corresponding exceedance probability and its associated benefit to corresponding accepted values. This could be desired, in particular, for hazards whose potential large consequence, catastrophic accidents are of primary concern. One would then evaluate the new risk by comparison of its exceedance probability (for any given loss) and benefit point to a curve of past accepted probabilities of exceeding such losses versus benefit in the same way as for expected loss and benefit comparisons.

In the foregoing procedure uncertainties enter into the new hazard's risk and benefit estimates. It is assumed here that the uncertainties in the preference curves can be neglected due to their derivation from a large sample of past experience or because, if they are subjectively established from expressed preferences, uncertainties are already accounted for in these preferences. If this assumption is not justified, then the curves need only be adjusted to represent, instead of point estimates of past or expressed risk-benefit preferences, bounds on these estimates: lower bounds on risk and upper bounds on benefit. For the new risk estimate, in whatever terms it is expressed, the presence of uncertainty implies that a confidence interval needs to be considered, and the upper confidence bound for some reasonable confidence level employed as the risk estimate in the comparison. Similarly, a lower confidence bound on the benefit would be developed and used.

The result would then be as shown in Figure 8-4. The new activity's expected loss versus benefit (a) is in the acceptable range, since its expected loss versus benefit point falls below the corresponding preference curve (1). But its exceedance probability versus benefit (b) does not fall below the corresponding preference curve (2), and so its catastrophic potential results in the hazard's non-acceptability.

As also shown in Figure 8-4, uncertainties in the new activity's estimates would cause the two points to move up (reflecting upper confidence bounds on the risk measures) and to the left (reflecting a lower bound on the benefit estimate) and so increase the potential for the new activity's risk-benefit points falling above their respective preference curves.

8.4 UNCERTAINTIES IN RISK-COST-BENEFIT EVALUATIONS

The treatment of uncertainties in several risk-cost-benefit evaluation procedures that have been employed is next examined. These procedures are:

- Equi-benefit comparisons
- Balancing of risks and benefits (including utility theory applications)
- Cost-effectiveness evaluations

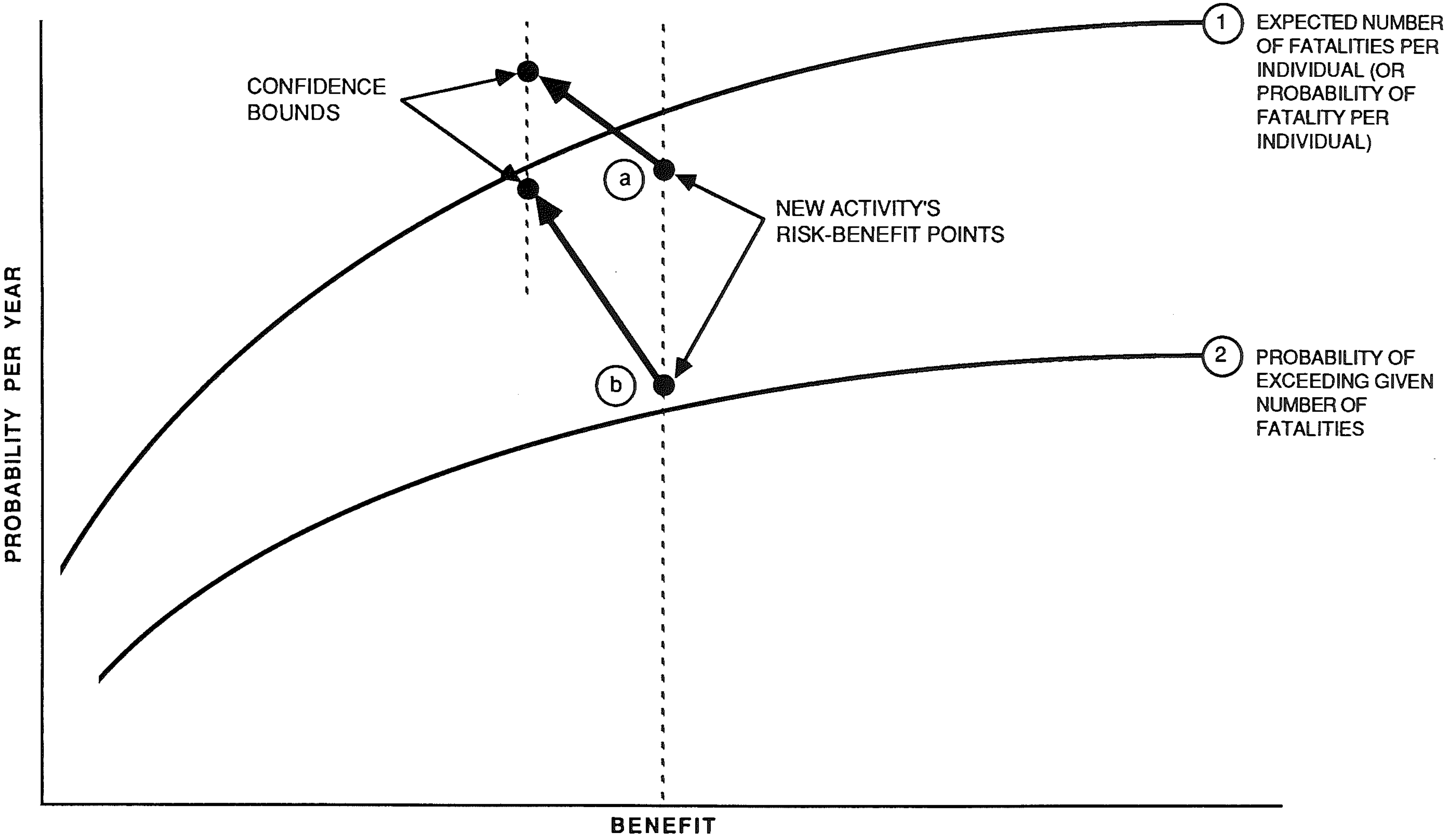

CONFIDENCE BOUNDS
a
b
NEW ACTIVITY'S RISK-BENEFIT POINTS
1
EXPECTED NUMBER OF FATALITIES PER INDIVIDUAL (OR PROBABILITY OF FATALITY PER INDIVIDUAL)
2
PROBABILITY OF EXCEEDING GIVEN NUMBER OF FATALITIES
PROBABILITY PER YEAR
BENEFIT

FIGURE 8-4. COMPARISON OF NEW HAZARDOUS ACTIVITY'S RISK AND BENEFIT ESTIMATES TO PREFERENCE CURVES (1) AND/OR (2) (e.g., INVOLUNTARY RISK-TAKING CASE)

8.4.1 Equi-benefit Comparisons

This evaluation procedure considers alternative activities for providing a common set of benefits, and determines the activity which does this at lowest risk. (The activity determined to be preferred in this way may, of course, change as the conditions assumed in the evaluation vary over their possible ranges.)

The measure of risk that has so far been employed in many applications of this procedure is expected loss; in particular, expected fatalities. A "risk accounting" is carried out, as feasible, for all subactivities involved in the provision of the ultimate benefits. The set of such subactivities that should be included in the accounting has been one aspect of the controversy that has arisen on the proper use of equi-benefit evaluations (e.g., should traffic accidents while they are going to work of additional coal miners required to produce a given amount of energy be considered? Accidents of truck drivers carrying food to stores serving the miners?) Another aspect of the controversy is whether or not fatalities that would result if the activity's workers were in some other occupation should in some way be subtracted from the activity's fatalities that are accounted.

The practical limits on such considerations are well beyond the scope of the present handbook. Rather, it will only be assumed that whatever these limits are, the uncertainty in them merely adds to the uncertainty in the overall risk estimate for the activity. A simple way to handle this addition is to consider an "Other" category of risk sources in conjunction with all the sources of the sub-activities considered. Then, in the comparative evaluation, each alternative activity's risk estimate is assigned a range of values from a minimum to a maximum for a given level of confidence. The decision on which alternative is to be preferred is then straightforward if one has a maximum value lower than all the other alternatives' minimum values. If no one alternative has this property, judgment must be employed as to whether one alternative can still be selected as the best because, while not uniformly dominant, its risk range nevertheless appears most favorable relative to the others. Figure 8-5 illustrates the situation. Alternative C is preferred straightforwardly in this case.

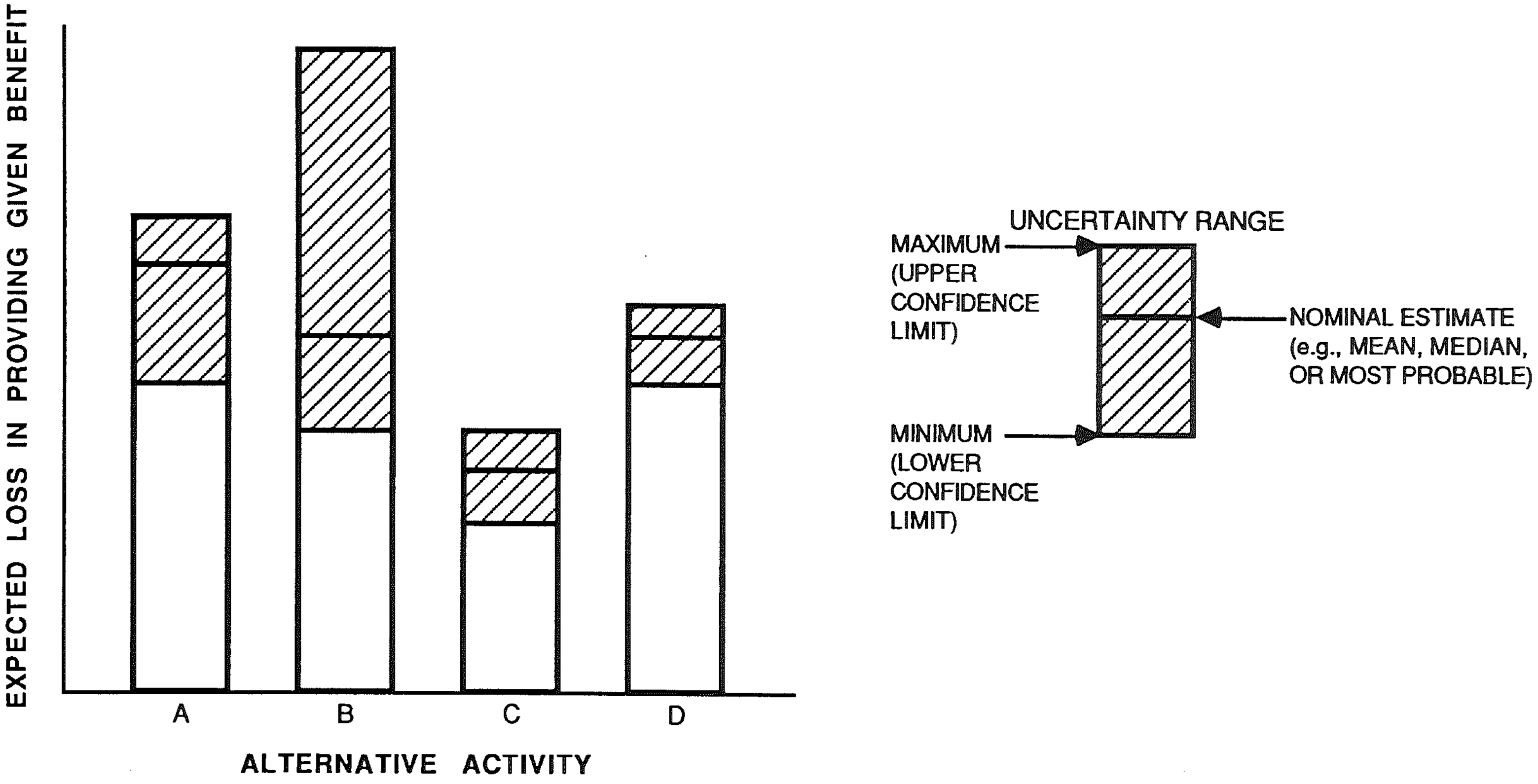


FIGURE 8-5. ILLUSTRATIVE EQUI-BENEFIT RISK COMPARISONS CONSIDERING UNCERTAINTIES

Refinements of this procedure are possible. For instance, the expected losses of each alternative activity's subactivities will generally accrue from different population classes, e.g., the bystanding public and workers. A relative weighting factor (reflecting the involuntary versus voluntary risk acceptance characteristics of the different classes) could be assigned to the different losses and expected weighted losses accumulated for the final comparison. A dominant alternative might then become evident, such as one that is more hazardous to workers but less to the public, but due to the weighting factors its maximum expected weighted loss is lower than the minimum expected weighted losses of all the alternatives. Another possibility, accomplishing the same result, would be to weight the alternatives' losses by the alternatives' economic costs, so that a lower-cost activity with a lower risk-times-cost factor after taking uncertainties into account would be preferred. Finally, the use of weights for higher losses in calculating the expected loss values (such as raising the losses to some power greater than one, as has previously been noted) may be desirable in order to disfavor alternatives with greater potentials for larger losses. Clearly, considerable further investigation would be needed for such procedures to become usable and acceptable.

8.4.2 Balancing of Risks and Benefits

All of the considerations in the preceding section also apply when the alternative activities are not first sized to provide the same benefits. In addition, however, the uncertainties that arise in expressing the risks and benefits in common, usually economic, terms must now be considered. The establishment of a "value-of-a-life" or related economic measure is a key contributor to these additional uncertainties.

In its simplest application, risk-benefit balancing entails accepting a hazardous activity if the evaluation of the risk and benefit in common terms indicates that its benefit exceeds its risk. More generally, the risk can be subtracted from the benefit for each of several alternatives, and the alternative with the highest resulting difference is then preferred. Direct cost could also be introduced into this procedure and subtracted from benefit for each alternative, if it is in the same terms as risk and benefit.

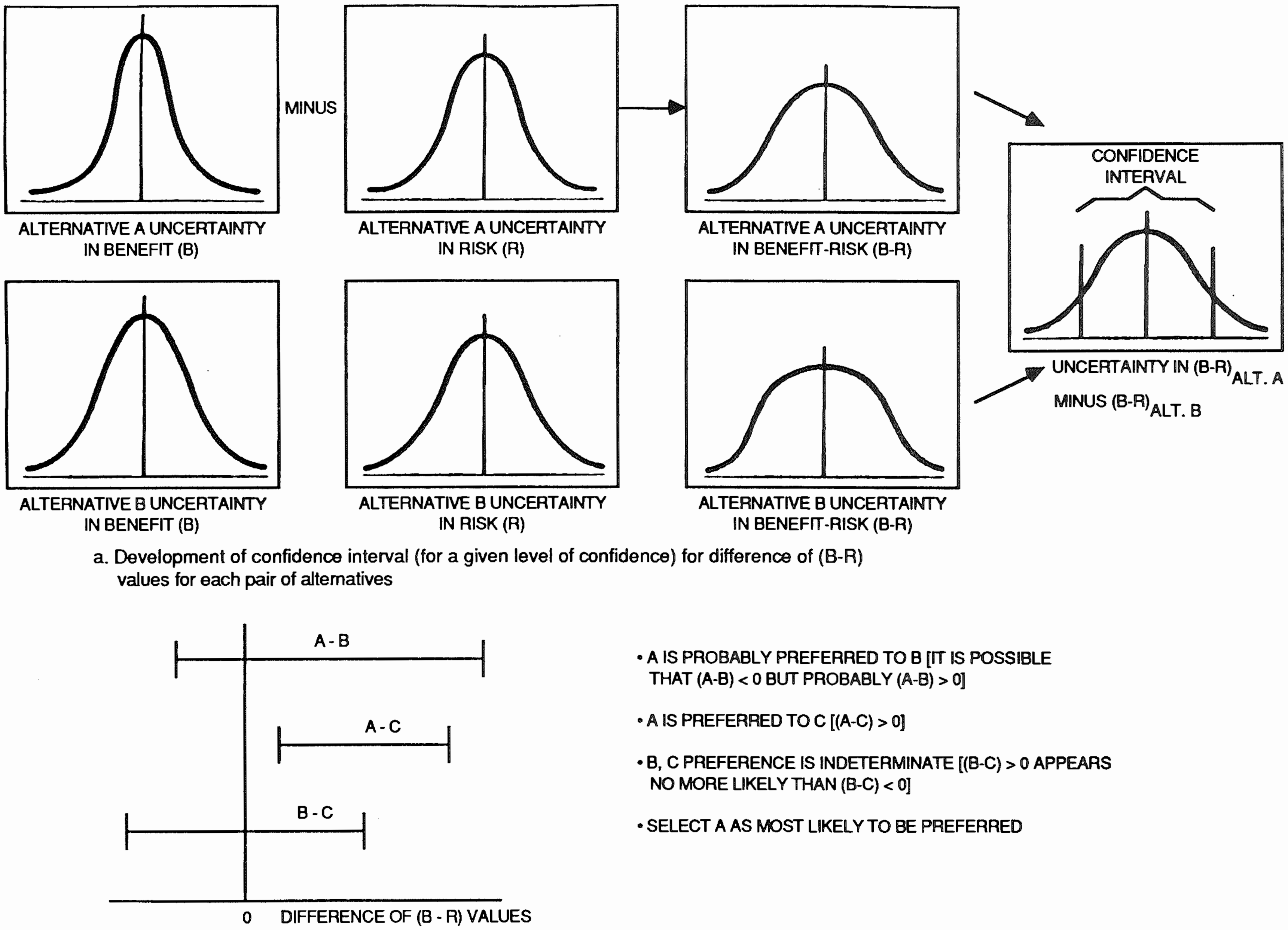
In treating the uncertainties in these decision processes, various techniques could be applicable. As in the preceding section, simple upper and lower bounds could be determined for the risks and for the benefits to reflect uncertainties in, for example, the value-of-a-life employed, and maximum risk compared to minimum benefit, in the activity acceptability evaluation case. Or, for the case involving a comparison of alternative activities, maximum benefit-minus-risk values could be considered. A more sophisticated technique would be to consider the uncertainty distributions of the risk and benefit estimates and establish the distribution of their difference for each alternative activity. This would be easy to do, in particular, if both could be assumed to be normally distributed so that their difference would be too.*

The distribution of the difference of such differences for each pair of alternative activities would then also be normal, and confidence intervals for a given level of confidence for these pair-differences would then aid deciding which alternative in each pair would be preferred. Doing this for all possible pairs would aid the selection of a preferred alternative overall.** Figure 8-6 illustrates this procedure. Note that, depending on the extent of the uncertainties present, it may or may not lead to a fully determinate preference, but in any case it would aid the judgmental selection of one.

It finally may briefly be noted that instead of a common evaluation of risks, costs, and benefits in economic terms, with the problems that have been noted, it is in principle possible to assess them in terms of measures of utility that incorporate a decision-maker's view of the relative importances of these factors, over all their dimensions and ranges of values. Now an expected utility measure is determined for each alternative activity which combines all of the activity's positive and negative attributes and which incorporates the probabilities of all their possible levels. The preferred alternative is then the one with the highest expected utility. Since the procedure is fundamentally subjective (but objective data and estimates can be used to aid judgment where they are available), it incorporates implicitly the decision-maker's uncertainties in these probabilities and so no explicit treatment of them is required. It need only be noted that the uncertainties will be reduced, at least in principle, if the decision-maker is provided improved risk, cost, and benefit information on the activities he is evaluating.

* If normality cannot be assumed, Monte Carlo procedures for calculating the convolution integral giving the difference distribution would still be possible, albeit with increased computational requirements.

** Alternatively, hypothesis testing could be employed.


MINUS
ALTERNATIVE A UNCERTAINTY IN BENEFIT (B)
ALTERNATIVE A UNCERTAINTY IN RISK (R)
ALTERNATIVE A UNCERTAINTY IN BENEFIT-RISK (B-R)
ALTERNATIVE B UNCERTAINTY IN BENEFIT (B)
ALTERNATIVE B UNCERTAINTY IN RISK (R)
ALTERNATIVE B UNCERTAINTY IN BENEFIT-RISK (B-R)
CONFIDENCE INTERVAL
UNCERTAINTY IN $(B\text{-}R)_{ALT.\ A}$ MINUS $(B\text{-}R)_{ALT.\ B}$
A - B
A - C
B - C
0
DIFFERENCE OF (B - R) VALUES
• A IS PROBABLY PREFERRED TO B [IT IS POSSIBLE THAT (A-B) < 0 BUT PROBABLY (A-B) > 0]
• A IS PREFERRED TO C [(A-C) > 0]
• B, C PREFERENCE IS INDETERMINATE [(B-C) > 0 APPEARS NO MORE LIKELY THAN (B-C) < 0]
• SELECT A AS MOST LIKELY TO BE PREFERRED

a. Development of confidence interval (for a given level of confidence) for difference of (B-R) values for each pair of alternatives

b. Comparison of pairs' confidence intervals

FIGURE 8-6. ILLUSTRATION OF COMPARISON OF BENEFIT-RISK (B-R) OF ALTERNATIVE ACTIVITIES, INCORPORATING UNCERTAINTIES

8.4.3 Cost-Effectiveness Evaluations

This procedure is the least arbitrary of all those that have been considered in this handbook. It is limited here to selecting an activity from a set of alternatives that either provides the least risk for a given cost, or attains a required risk decrease at the lowest cost. Like the equi-benefit comparison procedure, it therefore does not need risk, cost, or benefit in common terms and so avoids such issues as that of the value-of-a-life. However, it retains any problems of comparisons over several attributes if risk or cost have more than one dimension.

The most common form of application of the cost-effectiveness decision support procedure in the context of safety is that of the comparison of alternative risk mitigating measures. In this case, effectiveness is the predicted decrease in risk (the "Δ risk") from some baseline activity's configuration that would result from the introduction of each alternative mitigation. The alternative that provides the maximum Δ risk for a given allowable cost budget, or that provides a required Δ risk at the minimum cost, is then preferred.

Uncertainties can arise in both the Δrisk prediction and the cost estimate. In the Δ risk maximization procedure, a conservative approach to handling the cost uncertainty is to consider only alternatives whose upper bounds, for a given level of confidence, remain under the allowable budget. For these alternatives, the one with the highest Δrisk lower bound, for a given level of confidence, would then be preferred. Figure 8-7 illustrates this simple process. Alternative A is the preferred choice. The analog for the cost minimization procedure is obvious.

There are less conservative approaches that could also be followed, including, for example, obtaining the uncertainty distributions for the alternatives' risk estimates and then developing the joint distributions of their differences in pairs (in a way similar to that illustrated in Figure 8-6). Assessing the confidence intervals for the difference pairs enables the identification of the alternative that, at the given confidence level, is most likely to be superior to all of the others.

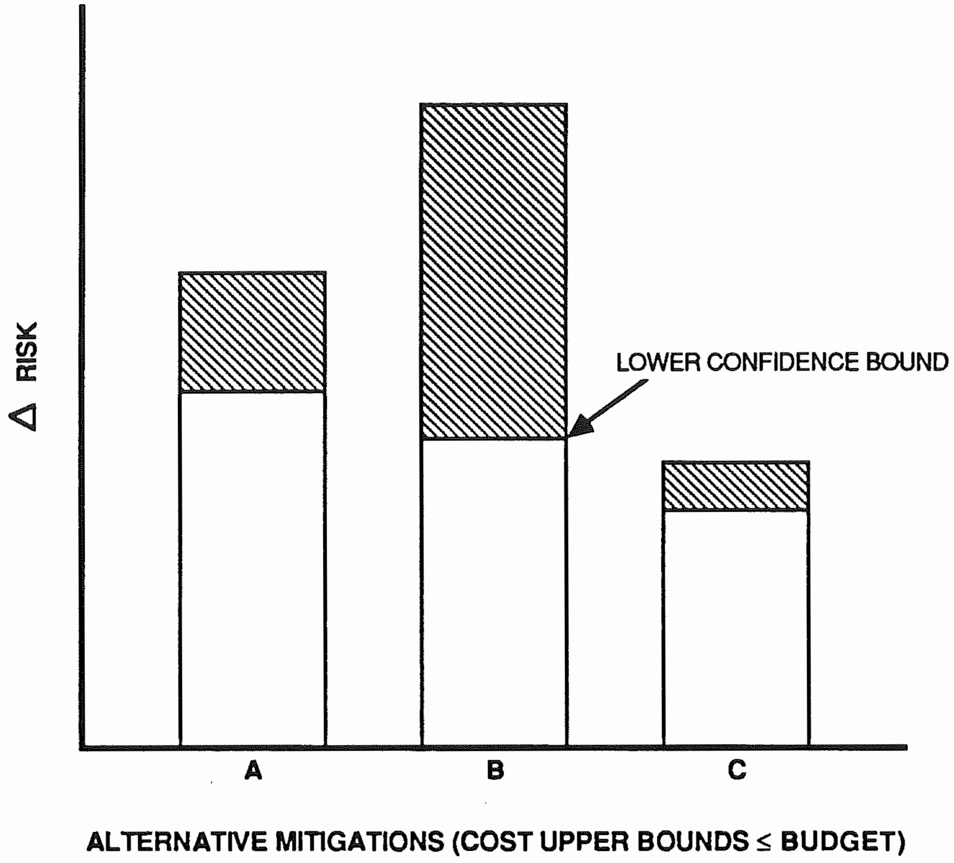


FIGURE 8-7. ILLUSTRATIVE Δ RISK-MAXIMIZATION COST-EFFECTIVENESS COMPARISON PROCEDURE

The treatment of uncertainty when only single-attribute comparisons need to be made is thus not an especially difficult problem. If, however, several attributes are involved, as with multiple risk components (fatalities, injuries, etc.), significant difficulties arise. They are most easily resolved by reducing the comparison again to that of values of a single attribute by establishing relative importance weights for the different original attributes (decreases in the several risk components) and developing the weighted sum of their levels. The preferred alternative is then the one with the largest weighted risk decrease (assuming all alternatives' costs are equal).

With uncertainties present, if the uncertainty in each component risk decrease is normally distributed, so is their weighted sum, and so is the difference of such sums for each pair of alternatives. (If normality cannot be assumed, a Monte Carlo procedure can again be employed, at least in principle.) Once again, the procedure illustrated in Figure 8-6 aids identification of the preferred alternative on the basis of comparisons of the confidence intervals for a given confidence level on the differences for all pairs.

Finally, instead of the simple subjective importance weights, the more controllable procedures of utility theory can arrive at a similar end, a single value, now the expected utility, for each alternative, on the basis of which the preferred alternative is determined. As in the preceding subsection, it is assumed here that the utility function enabling this procedure already incorporates all relevant uncertainties.

8.5 FINAL REMARKS

The present chapter has attempted to review in an organized manner the impacts on risk decision-making procedures of uncertainties in the risk estimates (and in associated estimates of benefits and costs). It has attempted to bring out the main considerations in these impacts and, in an admittedly rudimentary manner, describe possible ways in which they can be treated in risk acceptance and/or risk-cost-benefit decision processes. The initial descriptions of the treatments given here should provide bases for further investigations of more complete means for handling the uncertainties appropriately to their definitions in any given decision problem.

It is clear that even in the simplest forms of risk evaluation (ambient risk, equi-benefit and cost-effectiveness comparisons, with only one risk dimension) in risk decision-making, judgment enters in importantly in the treatment of the uncertainties in the associated decision process. A maximum credible uncertainty range, or, with a somewhat more sophisticated approach, a reasonable confidence level, must be selected. For the more complex procedures, the presence of uncertainties may make a fully determined decision impossible, and judgment is required to compare the impacts of the uncertainties (reflected in confidence intervals) and the preferences among the various alternatives. When the risk is considered in several dimensions, either inherently subjective relative importance weights or utility functions are required implicitly.

CHAPTER 9

RECOMMENDED QUALITATIVE AND QUANTITATIVE RISK ASSESSMENT AND RISK DISPOSITION DECISION PROCEDURES

The concepts and methods that have been delineated in the previous sections of this manual are now placed into the perspective of an overall risk assessment and risk disposition process for NASA. A recommended approach to a qualitative process is first described, and, then, in cases where its employment is appropriate, the recommended approach is extended to one for a full quantitative process. It is emphasized that these approaches are intended to be generally applicable, but they can be expected to require tailoring in each particular case, both in the factors considered and in the way in which they are evaluated. Note also that in complex programs both qualitative and quantitative risk assessments may be conducted for different areas of the program. Their integration to support top level decision-making then must be considered.

9.1 QUALITATIVE RISK ASSESSMENT AND RISK DECISION-MAKING

The qualitative assessment of the risks of mishaps for a given program or facility element should be based on the "5x5" hazard frequency/severity categorization matrix shown in Figure 9-1. The assignment of matrix cells to mishaps should be derived with a qualitative fault tree model, supported by basic hazards analyses and failure modes and effects analyses.

9.1.1 Fault Tree Modeling

As discussed in Chapters 3 and 4, each identified mishap of concern should be represented as a Top Event (TE) of a fault tree or equivalent model. The tree should be developed down to the level of events (basic or primary events) for which judgments on occurrence frequency categories can most reasonably be made. The selection of lowest level events (LLEs) involves the tradeoff between increased fault tree size and complexity, and generally greater ease and precision in the frequency judgments, as lower levels are considered. A relative frequency category should be assigned to each LLE, based on a five

SEVERITY

FREQUENCY OR LIKELIHOOD OF OCCURRENCE

	5	4	3	2	1
5	■	▩	▩	▧	▧
4	▩	▩	▩	▧	▧
3	▩	▩	▨	▧	▧
2	▨	▨	▧	□	□
1	▨	▨	▧	□	□

■	THE MOST SEVERE AND MOST LIKELY TO OCCUR OF ALL RELEVANT HAZARDS' POTENTIAL CONSEQUENCES. MANAGEMENT ACTION AND CORRECTIVE ACTION ARE MANDATORY (CERTAINLY UNACCEPTABLE)
▩	CRITICAL IN SEVERITY AND LIKELIHOOD OF OCCURRENCE SUFFICIENT TO WARRANT MANAGEMENT ATTENTION AND CORRECTIVE ACTION (PROBABLY UNACCEPTABLE)
▨	SERIOUS ENOUGH IN SEVERITY OF OCCURRENCE TO INFORM MANAGEMENT AND PROBABLY REQUIRE CORRECTIVE ACTION (FOR ANY LIKELIHOOD OF OCCURRENCE) (POSSIBLY ACCEPTABLE OR UNACCEPTABLE)
▧	NOT AS SEVERE OR NOT AS LIKELY TO OCCUR BUT SHOULD BE PRESENTED TO MANAGEMENT FOR INFORMATION AND POSSIBLE CORRECTIVE ACTION (PROBABLY ACCEPTABLE)
□	SEVERITY AND LIKELIHOOD OF OCCURRENCE ARE BOTH LOW ENOUGH TO REQUIRE PRESENTATION TO MANAGEMENT FOR INFORMATION ONLY (CERTAINLY ACCEPTABLE)

FIGURE 9-1. ILLUSTRATIVE HAZARD FREQUENCY/SEVERITY CATEGORIZATION MATRIX

point scale (1 for lowest frequency, 5 for highest), in accordance with Table 9-1 and Figure 9-2 a through f. The severity category should be assigned to each TE (i.e., identified mishap), based on a five point scale (1 for lowest severity, 5 for highest), in accordance with Table 9-1 B.

9.1.2 Mishap Modes Delineation

The various modes of occurrence of each TE should be delineated as the mincutsets (see Chapter 4 and Fault Tree Handbook, Vesely, 1981) of the fault tree for that TE. Each mincutset or mishap mode consists of one or more LLEs such that if they all occur, the mishap will occur. Thus, the (essentially) simultaneous occurrence of all the LLEs in a mincutset is one mode by which the mishap will occur. The total ensemble of mincutsets or modes to be considered for a TE or mishap may be reduced by deleting all mincutsets with more than a specified number of LLEs, under the presumption (which should be confirmed by judgmental analysis) that more than the specified number of LLEs can occur simultaneously only with negligibly small frequency, and so such "higher order" mincutsets cannot contribute significantly to the frequency of occurrence of the mishap.

The TE is then the logical union of its (remaining) mincutsets and its frequency of occurrence is approximately the sum of the frequencies of occurrence of its mincutsets. (Refinement of this approximation may be necessary in cases involving relatively higher frequency mincutsets or especially large numbers of mincutsets.) Each mincutset is the logical product of its LLEs. The frequency of occurrence of a mincutset is approximately the product of the frequencies of occurrence of its LLEs. (This approximation also may require refinement in cases in which non-independent LLEs are involved, e.g., due to common causes of the occurrence of several LLEs in a mincutset.) The proper treatment of these sums and products of qualitative frequencies is described in Section 9.1.4, below.

9.1.3 Component Importance Analysis

The qualitative importance of an LLE to the occurrence of the TE can be appraised with several methods. In essence, and with various refinements, they rank the LLEs in a fault tree according to the number of times they occur in the mincutsets that cause the TE. In some methods this is done also taking into account the numbers of other events also included in the mincutsets, weighting a LLE's importance higher if it is in lower multiplicity mincutsets than if it is in higher ones. For example, write:

A. Frequency (Reference: Related Document 1.4f.)

5-
4-
3- Weighted sum of subparameter categories: see Figures 9a-9e, 9f.*
2-
1-

B. Severity (Reference: Related Document 1.4f. with modifications)

5- Loss of personnel or system

4- Major injuries to personnel or damage to system

3- Loss of mission, minor injuries to personnel, or minor damage to system

2- Loss of major mission objective

1- Loss of minor mission objective

* NOTE: The particular values of subparameter scores and weights shown are subject to change for each program or facility.

TABLE 9-1. FREQUENCY AND SEVERITY CATEGORIES

SUBPARAMETER	VALUE DEFINITION	VALUE
A. CAUSE(S)	CAUSED BY A COMBINATION OF TWO OR MORE EVENT/FAILURES WHERE EACH EVENT/FAILURE OCCURRENCE IS INDEPENDENT OF THE OTHERS	1.
	CAUSED BY TWO OR MORE REDUNDANT ELEMENT FAILURES OR EVENTS	2.
	CAUSED BY TWO OR MORE DEPENDENT ELEMENT FAILURES OR EVENTS	3.
	CAUSED BY A PARTICULAR SINGLE ELEMENT FAILURE OR EVENT, OR BY PERSONNEL ERROR	4.
	CAN OCCUR FROM ANY ONE OF SEVERAL CAUSES	5.
	VALUE CHOSEN	
	TIMES WEIGHTING FACTOR	x 4 =
	WEIGHTED VALUE	

FIGURE 9-2a. OCCURRENCE FREQUENCY SUBPARAMETER (CAUSE) EVALUATION

SUBPARAMETER	VALUE DEFINITION	VALUE
B. CONTROLS	PRECLUDED FROM OCCURRENCE BY MORE THAN ONE INDEPENDENT CONTROL MEASURE, AT LEAST ONE OF WHICH IS A HARDWARE CONTROL	1.
	PRECLUDED FROM OCCURRENCE BY ONE HARDWARE CONTROL	2.
	A) PRECLUDED FROM OCCURRENCE BY AN AUTOMATICALLY OPERATED SAFETY DEVICE OR SAFETY SYSTEM B) ACCEPTED BECAUSE OF TESTED DESIGN FACTORS, QUALITY CERTIFICATION TESTING, AND INSPECTIONS	3.
	PRECLUDED FROM OCCURRENCE BY A MANUALLY INITIATED PROCEDURE INSTIGATED BY A CAUTION OR WARNING SIGNAL	4.
	PRECLUDED BY A MANUALLY INITIATED PROCEDURE	5.
	VALUE CHOSEN	
	TIMES WEIGHTING FACTOR	x 4 =
	WEIGHTED VALUE	

FIGURE 9-2b. OCCURRENCE FREQUENCY SUBPARAMETER (CONTROLS) EVALUATION

SUBPARAMETER	VALUE DEFINITION	VALUE
C. FAILURE HISTORY AND SYSTEM MATURITY	THE SYSTEM IS VERY MATURE AND HAS A LOW FAILURE RATE	1.
	THE SYSTEM IS MATURE AND HAS SOME FAILURES	2.
	THE SYSTEM IS NEW AND HAS NO FAILURE HISTORY OR THE SYSTEM IS YOUNG AND HAS A LOW FAILURE RATE	3.
	THE SYSTEM IS VERY MATURE AND HAS A HIGH FAILURE RATE	4.
	THE SYSTEM IS YOUNG AND HAS A HIGH FAILURE RATE	5.
	VALUE CHOSEN	
	TIMES WEIGHTING FACTOR	x 3 =
	WEIGHTED VALUE	

FIGURE 9-2c. OCCURRENCE FREQUENCY SUBPARAMETER (FAILURE HISTORY AND SYSTEM MATURITY) EVALUATION

SUBPARAMETER	VALUE DEFINITION	VALUE
D. METHODS OF DETECTING HAZARDOUS CONDITION	HAZARDOUS CONDITION CAN BE DETECTED BY TWO OR MORE INDEPENDENT MEANS	1.
	HAZARDOUS CONDITION CAN BE DETECTED BY A SINGLE INSTRUMENT	2.
	HAZARDOUS CONDITION CAN BE DETECTED BY DIRECT SENSORY OBSERVATION BY PERSONNEL	4.
	HAZARDOUS CONDITION IS NOT DETECTABLE	5.
	VALUE CHOSEN	
	TIMES WEIGHTING FACTOR	x 1 =
	WEIGHTED VALUE	

FIGURE 9-2d. OCCURRENCE FREQUENCY SUBPARAMETER (METHODS OF DETECTING HAZARDOUS CONDITION) EVALUATION

SUBPARAMETER	VALUE DEFINITION	VALUE
E. TIME TO EFFECT AMOUNT OF TIME FROM WHEN THE HAZARDOUS CONDITION BEGINS TO THE TIME WHEN CORRECTIVE ACTION IS NO LONGER POSSIBLE	ONSET RATE IS SLOW ENOUGH TO ALLOW SAFE TERMINATION OF MISSION OR OPERATION.	1.
	ONSET RATE IS FAST ENOUGH TO MAKE SAFE TERMINATION QUESTIONABLE.	3.
	ONSET RATE IS FAST ENOUGH TO MAKE SAFE TERMINATION IMPOSSIBLE.	5.
	VALUE CHOSEN	
	TIMES WEIGHTING FACTOR	x 1 =
	WEIGHTED VALUE	

FIGURE 9-2e. OCCURRENCE FREQUENCY SUBPARAMETER (TIME TO EFFECT) EVALUATION

SUM OF WEIGHTED VALUES FOR SUBPARAMETERS A THROUGH E =	
TOTAL VALUE DIVIDED BY 13 =	
WEIGHTED AVERAGE VALUE ROUNDED OFF TO NEAREST INTEGER (LESS THAN 0.5 - ROUND TO LOWER INTEGER) (0.5 OR GREATER - ROUND TO HIGHER INTEGER)	FINAL OCCURRENCE FREQUENCY CATEGORY

FIGURE 9-2f. FINAL OCCURRENCE FREQUENCY CATEGORY EVALUATION

Importance of LLE$_i$ = (Sum of the probabilities of all mincutsets assuming that the probability of LLEi = 1.0 and the probabilities of all other LLE's is 0.5) - (Sum assuming that the probability of LLEi = 0.0 and the probabilities of all other LLEs is 0.5)

The importance ranking then aids the prioritization of the components for design or other reliability enhancements. Quantitative importance analysis, discussed earlier in Section 4, improves this process.

9.1.4 Mishap and Mishap Mode Risk Categorization

Since the frequencies of the LLEs are established only categorically rather than numerically (as in the quantitative process to be described below), it will be necessary to employ a judgment-based procedure to derive the mincutset (mishap mode) and TE (mishap) frequency categories. Clearly, if all LLEs in a mincutset have category 1 (lowest) frequencies so must the mincutset. More generally, a mincutset's frequency category cannot be greater than the lowest frequency category among the LLEs composing the mincutset. If at least one mincutset causing a TE has category 5 (highest) frequency so must the TE. Again, more generally, a TE's frequency category cannot be less than the highest frequency category among the mincutsets causing the TE. In most cases, however, the mincutset or TE frequency category must be gauged from an appraisal of the frequency categories of the specific LLEs constituting the mincutset and the frequency categories of the specific mincutsets causing the TE. Whether some more structured procedures for doing this (e.g., a weighted geometric mean (rounded to an integer) of LLE categories for the mincutset, and a weighted mean (rounded to an integer) of mincutset categories for the TE) would be helpful and not over-arbitrary, should be examined further in each case.

Each mishap mode and each mishap now falls into a cell of the frequency/severity or risk categorization matrix shown in Figure 9-1. (The severity category for a mishap mode or mincutset is of course the same as the severity category for the mishap or TE it causes.) Two basic applications of the results can be made, as appropriate. These are next described.

9.1.5 Risk Acceptability Evaluation and Decision-Making

The risk accruing to a given mishap mode or to a given mishap for a given program or facility element, may be deemed acceptable, or require management attention and correction, depending on the cell its frequency and severity categories fall into in a matrix such as that shown in Figure 9-1. If the cell is in the lower right corner area, the risk may be deemed acceptable. If it is in the upper left corner area, the risk requires correction. If it is in the middle area (due, perhaps, to uncertainties) the case requires further analysis to make a decision on acceptance or mitigation. The uncertainties in the categorization could be analyzed with greater refinement, possibly including quantitative methods as noted later, to determine if a categorization into one of the corner areas could be justified. The analysis could incorporate assessment of the cost and risk improvement (i.e., frequency and/or severity category decrease) of potential corrective actions as discussed below.

The overall risk accruing to a program or facility element from all of its mishaps that have been considered can also be examined. Identify all mishaps or TEs of the element which have categorized severities of x or higher, where x = 1, 2, ..., 5. The frequency category of a mishap with at least severity category x is then judged from the frequency categories of the identified set of TEs with much the same approach as in judging a TE frequency category from its mincutsets' categories (Section 9.1.3). Some simplification in this results, however, from the fact that the TEs in an identified set for a given x are mutually exclusive, so that their true frequencies (but not directly their frequency categories) simply add to give the true frequency of occurrence of some mishap with severity at least x.

In summary, a basis can be established for deriving and assessing the acceptability of the risk associated with:

- Each mishap mode for a given program or facility element and mishap, from its frequency and severity categories
- Each mishap for a given element, from its frequency and severity categories
- Each element, from the frequency category for the occurrence of all possible mishaps whose severity categories are at least a given category.

9.1.6 Risk Mitigation Evaluation and Decision-Making

The second application of the frequency/severity or risk categorization of each mishap mode and each mishap as in the matrix of Figure 9-1 is that of the evaluation of alternative risk mitigations, such as design changes. Only the relative risk decreases provided by such mitigations are involved in this evaluation here. A complete evaluation would also incorporate considerations of each mitigation's cost, schedule, and performance impacts, and, possibly, the mitigation's potential for increasing risk in some other program or facility element.

Each alternative mitigation will affect the frequency categorizations of some particular set of LLEs in some particular manner. The effects of any changes in the LLE categorizations will propagate to changes in the categorizations of the mincutsets or mishap modes. These will propagate in turn to the categorizations of the TEs or mishaps, and finally of the sets of mishaps for the element whose severity categories exceed any given category. Thus, comparisons of risk decreases accruing to the alternative mitigations can be made at each level (mishap mode, mishap, program or facility element) and mitigations identified that provide the most improvement in frequency category.

Of course, the possible variations among the five categories are limited, and so it will probably often occur that more than one mitigation will provide the same categorized risk decrease and thus not be able to be discriminated on the basis of risk decrease alone. The other evaluation factors (cost, schedule, etc.) would then enter into the decision to select one from among several alternative mitigations all of which are roughly equally effective in their risk impacts.

The benefits of refining this decision process with the use of a full range of possible numerical values of assessed risks and risk decreases provide one important motivation for the quantitative risk assessment process next discussed.

9.2 QUANTITATIVE RISK ASSESSMENT AND RISK DECISION-MAKING

A quantitative fault tree model or an equivalent should be developed for each TE or mishap of concern whose qualitative categorization is too uncertain for risk management purposes, or when a refined risk acceptability analysis or risk mitigation evaluation is required.

9.2.1 Mishap Modes and Mishap Occurrence Probability Estimation

The quantitative fault tree extends the qualitative fault tree down to component (including sofware and human) failure events for which numerical failure rate/failure probability estimates can be made. These estimates should employ Bayesian techniques, as necessary, to make best use of engineering judgment as well as all applicable test and operating experience. The estimates should include not only point values for the rates or probabilities, but also uncertainty distributions or confidence intervals for the "true" values relative to the point values. These component point values and distributions can then be propagated up the fault tree to quantify the corresponding point values and uncertainty distributions for the mincutsets (mishap modes) and the TE (mishap).

9.2.2 Importance Analysis

Quantitative importance analyses of the relative significance to the TE's occurrence probability of particular components' failure probabilities should also be carried out to establish guidance for risk mitigation efforts, such as redesign of an important component, the addition of redundancy, etc.

9.2.3 Mitigations Evaluation

The effectiveness of such occurrence probability-reducing mitigations should be estimated in terms of the TE probability decreases they induce.

9.2.4 Risk Estimation, Risk Mitigations Evaluation and Decision-Making

Associated with the TE will also be the consequences and losses due to its occurrence. In some cases, mitigations may be available which are applicable to reducing the consequences (see Section 9.2.6, below). Their effectiveness shall be evaluated usually in terms of the decreases in the TE's expected consequences they induce. If a complete "risk profile" (complementary cumulative distribution of possible loss) is generated from all of a program or facility TEs' uncertainty distributions and associated losses (perhaps, also including uncertainty distributions for the losses themselves), a more comprehensive evaluation of candidate mitigations can be conducted in terms of the changes in the risk profile they induce.

Decision-making on risk mitigation iterates with decision-making on risk acceptability (or tolerability, when waivers are applicable) in the overall risk disposition decision process. The consideration of appropriate acceptability criteria is discussed below. One basis for acceptance or tolerance of a risk is the establishment of the infeasibility or extreme cost ineffectiveness of any means for its mitigation. However, more generally, candidate mitigations can be identified and evaluated in terms of their risk-decrease cost effectiveness. Standard procedures then apply to deciding on the best among them for implementation. These procedures include selecting the most effective (largest decrease in risk) mitigation among all candidates meeting a given cost limitation; the least costly mitigation among all candidates providing at least a given risk decrease; the mitigation providing the largest effectiveness-to-cost ratio. When several risk "dimensions" (i.e., different kinds of risk, such as risks to life, economic risks; and different targets of risk, such as voluntary risk-taking workers, involuntary risk taking population groups) must be considered, a more complex procedure is required. Utility theory techniques may be usable (see, e.g., Philipson, 1982; Keeney, 1980; and textbooks on utility theory).

9.2.5 Risk Acceptability Decision-Making

The potential for a possible mishap may be deemed acceptable if the mishap's probability of occurrence or its possible consequences are low. Whereas in a qualitative analysis these two factors must be treated separately, in a quantitative analysis they may be integrated naturally in the probability distribution of possible loss which has the expected (average) loss as its mean.

In some cases, the development of the entire probability distribution will not be required. The mishap potential (or its associated risk) could then be decided to be acceptable if the expected loss, or a credible upper bound on it, is sufficiently low in relation to some criterion. This criterion would take into account such relevant factors as the importance of the hazardous activity or mission, whether the risk is acceptable to the individuals exposed to the hazards either voluntarily or involuntarily, and the feasibility and costs of any available mitigation measures that could reduce the risk. If the risk, i.e., expected loss, is not sufficiently low to be clearly acceptable, but no feasible mitigation is available and yet the activity or mission must go forward, it may be decided to "tolerate" the risk by instituting a waiver or deviation. The quantitative analysis then provides a specific measure of the risk that is associated with such an action.

In the case that a complete probability distribution for possible loss is developed, it becomes possible to assess not only the acceptability of the expected loss but also the acceptability of the chances of occurrence of different ranges of possible loss. Thus, a relatively high probability of small values of loss may be acceptable, but the probability of large values of loss may not be low enough for the "tail of the curve" to be acceptable. Conversely, if the tail of the curve is acceptable, it may yet be the case that the probability of smaller losses of some significance is too high. Consideration of only expected loss, or worst-case loss, does not enable this type of more complete assessment.

9.2.6 Consequence Analysis and Loss Estimates

This area of analysis of the risks for a specific program or facility often most differs from that in analyses for other hazardous activities. It extends from the basic engineering analyses that led to the designs of the various program or facility elements with which hazards of significance are associated. It develops and integrates models of the immediate effects of each particular mishap of concern, of the propagation of these effects through the activities of the program or facility, and of the losses that then result. These models may be deterministic and produce only expected or worst case estimates of loss given the occurrence of a particular mishap, or they may need to be probabilistic, taking uncertainties explicitly into account, and produce (conditional) probability distributions of the possible losses given that the mishap occurs. These estimates or distributions then combine with the mishap occurrence probability distribution to produce the (marginal) loss probability, or risk, distribution upon which risk disposition decision-making should be based.

9.3 INTEGRATION OF QUALITATIVE AND QUANTITATIVE RISK ASSESSMENTS

For complex systems or facilities for which qualitative risk assessments may be made for some elements and quantitative assessments for others, it may be necessary that such mixed assessments of lower level elements' risks can be consistently combined to produce assessments of risks at higher levels. As described in Section 9.2, well established techniques exist for doing this when all risks are expressed quantitatively. Means for the consistent combination of all-qualitative (categorized) risks were discussed in Section 9.1, with judgment playing an important role in the combination just as it does in the assessments of the individual risks. Procedures for combining qualitative with quantitative assessments of risks to enable integrated assessments of higher level risks require development. An initial approach is described in this section.

Fundamentally, what is needed is either a categorization of numerically assessed risks consistent with the categorization of the qualitatively assessed risks, or a numerical scaling of the qualitatively assessed risks consistent with the estimates of the numerically assessed risks. With the first approach, information about the numerically assessed risks is discarded as their description becomes less precise. With the second approach, information is added about the qualitatively assessed risks, information that can only be generated judgmentally. A utility analysis procedure could carry the second approach to its fullest development. However, its implementation could be burdensome and a simpler scaling procedure ought to be considered before a complete utility analysis is proposed.

Such a scaling procedure is next outlined.

The basic idea is to develop reasonable quantitative scaling of the qualitative risks by calibrating the qualitative hazard frequencies and severities through judgmental comparisons to frequencies and severities, or losses, that have been quantified for other hazards. A structured process is used to maintain consistency among the resulting quantifications and to minimize the necessary effort.

9.3.1 Frequency Estimates

The following steps are carried out:

Step 1. Judgmentally order the hazards' categorized frequencies in a sequence by decreasing magnitude: f_1, f_2, ... f_n, ... Ties are permitted. The ordering is to be done as well as possible. In principle, it could be no more refined that the original categorization, but the more refined it can be, the better.

Step 2. Judgmentally establish the relative magnitude, r_n, of f_{n+1} compared to f_n; i.e., $f_{n+1} = r_n f_n$, $0 < r_n \leq 1$. Again, this is to be done with as much refinement as possible, but the process applies even if a lack of information necessitates a number of ties (r's equal to 1). Note that the process requires only the ability to compare successive pairs of frequencies that generally should not differ greatly. Absolute judgments of numerical frequencies are avoided.

Step 3. Identify a quantified risk (elsewhere in the system, in a similar system, or a comparable generic risk) with a frequency, estimated as f, that can be judged to be approximately the same as the frequency, f_k, of one of the ordered qualitative frequencies. (See Step 6, below, if no comparable quantified risk can be identified.)

Step 4. Quantify all of the ordered qualitative frequencies, using

$$f_{k+1} = r_k f_k,\ f_{k+2} = r_{k+1} \cdot r_k f_k,\ \text{etc.} \tag{9-1}$$
$$f_{k-1} = (1/r_{k-1}) f_k,\ f_{k-2} = (1/r_{k-2}) \cdot (1/r_{k-1})\ f_k,\ \text{etc.}$$

Step 5. Check the results for consistency by comparing some resulting frequency values in the sequence with any other available quantified frequencies. Judgmentally adjust the sequence of values as necessary.

Step 6. If no comparable qualitative frequencies are available, Step 3 is replaced by a judgment of the quantitative value of one frequency it is most convenient to estimate absolutely; usually, this is the highest frequency, f_1, of the most likely to occur, and so easiest to comprehend, event.

It is evident that while reasonable and not unduly difficult to carry through, the accuracy of the foregoing process can suffer from the propagation and buildup of errors as the successive estimates are made. It is important, therefore, to verify the results wherever possible by comparisons with independent quantitative estimates of the frequencies of other similar hazards. The results should then be adjusted for maximum consistency.

9.3.2 Severity Estimates

A similar procedure is carried out for the severity components of the qualitative risk assessments. The categorized severities are ordered according to increasing potential loss magnitude. Except for this, Steps 1-6 above are executed, with "frequency" replaced by "severity or loss."

9.3.3 Integrated Risk Assessment

With all hazard frequencies and severities quantified, the numerical risk induced by each hazard and each subsystem's or system's set of hazards is finally delineated numerically in one of two ways. Risk as expected loss is the sum over the relevant hazards' of the products of the hazard frequencies and severities (losses). Risk as a risk profile is the curve of the frequency versus loss or the (complementary cumulative) curve of the frequency of exceeding each possible level of loss because of a hazard or set of hazards.

REFERENCES

Abernethy, R.B., Breneman, J.E., Medlin, C.H., and Reinman, G.L., *Weibull Analysis Handbook.* Prepared for the Aero Propulsion Laboratory, Wright-Patterson AFB, Ohio 45433, AFWAL-TR-83-2079, Pratt & Whitney Aircraft, West Palm Beach, Florida, November 1983.

Ahmed, S., D.R. Metcalf, R.E. Clark, and J.A. Jacobsen, *BURD--A Computer Program for Bayesian Updating of Reliability Data,* NPGD-TM-582, Babcock & Wilcox, 1981.

Air Force Rome Air Development Center, *Nonelectronic Reliability Notebook*, RADC-TR-75-22, AD/A005-657, January 1975.

Allen, D.J., and Rao, M.S., "New Algorithms for the Synthesis and Analysis of Fault Trees," *I & C Fundamentals,* Vol. 19, pp. 79-85, February 1980.

Apostolakis, G., *Mathematical Methods of Probabilistic Safety Analysis*, School of Engineering and Applied Science, UCLA Report UCLA-ENG-7464, 1974.

Apostolakis, G., and A. Mosleh, "Expert Opinion and Statistical Evidence: An Appliction to Reactor Core Melt Frequency," *Nuclear Science and Engineering*, Vol. 70, pp. 135-149, 1979.

Apostolakis, G., Kaplan, S., Garrick, B.J., and Duphily, R.J., "Data Specialization for Plant-Specific Risk Studies," *Nuclear Engineering and Design*, Vol. 56, pp. 321-329, 1980.

Apostolakis, G., and Kaplan, S., "Pitfalls in Risk Calculations," *Reliability Engineering*, Vol. 2, pp. 135-145, 1981.

Atwood, C.L., *Data Analysis Using the Binomial Failure Rate Common Cause Model.* Prepared for U.S. Nuclear Regulatory Commission, EG&G, Idaho, NUREG/CR-3437, EGG-2271, September 1983.

Bain, L.J., *Statistical Analysis of Reliability and Life-Testing Models*, Marcel Dekker, Inc., New York, 1978.

Basu, A.P. and El Mawaziny, A.H., "Estimates of Reliability of k-out-of-m Structures in the Independent Exponential Case," *J. Amer. Stat. Assn.*, Vol. 73, pp. 850-854, December 1978.

Birnbaum, Z.W., "On the Importance of Different Components in a Multicomponent System," in *Multivariate Analysis* - II, P.R. Krishnaiah (ed.), Academic Press, New York, 1969.

Blyth, C.R., "Approximate Binomial Confidence Limits," *J. Amer. Stat. Assn.*, Vol. 81, pp. 843-855, 1986.

Box, G.E.P. and Tiao, G.C., *Bayesian Inference in Statistical Analysis*, Addison-Wesley Publishing Co., Reading, Massachusetts, 1973.

Dougherty, E.M., Jr. and Fragola, J.R., "Foundations for a Time Reliability Correlation System to Quantify Human Reliability," *Proceedings IEEE Fourth Conference on Human Factors and Power Plants*, Monterey, California, June 1988.

Dunglinson, C. and Lambert, H.E., "Interval Reliability for Initiating and Enabling Events," *IEEE Transactions on Reliability*, Vol. R-32, June 1983.

Fairbanks, K., Madsen, R., and Dykstra, R., "A Confidence Interval for an Exponential Parameter from a Hybrid Life Test," *J. Amer. Stat. Assn.*, Vol. 77, pp. 137-140, 1982.

Fairley, W.B., "Criteria for Evaluating the 'Small' Probability of a Catastrophic Accident From the Marine Transportation of Liquefied Natural Gas," *Proceedings of the Engineering Foundation Conference on Risk-Benefit Methodology and Application*, Asilomar, California, September 1975.

Fiksel, J., "Toward a De Minimis Policy in Risk Regulation," *Risk Analysis*, Vol. 5, pp. 257-259, 1985.

Fischoff, B., Slovic, P., and Lichtenstein, S., "Lay Foibles and Expert Fables in Judgments About Risks," in R. O'Riordan and R.K. Turner (eds.), *Progress in Resource Management and Environmental Planning*, Vol. 3, John Wiley & Sons, Chichester, England, 1981.

Fleming, K.N., "A Reliability Model for Common Mode Failure in Redundant Safety Systems," *Proceedings of the Sixth Annual Pittsburgh Conference on Modeling and Simulation*, San Diego, CA., April 1975.

Fragola, J.R. et al., *Formulation and Use of Reliability Data in Probabilistic Safety Assessment.* Prepared for the International Atomic Energy Agency, Science Applications International Corporation, July 1987.

Fussell, J.B. et al., *A Collection of Methods for Reliability and Safety Engineering*, Aerojet Nuclear Co., ANCR-1273, April 1976.

Gateley, W.Y. et al., *The GO Modeling Manual*, Kaman Sciences Corporation, K80-36U, March 1980.

Hall R.E., Fragola, J.R., and Wreathall, J., *Post Event Human Decision Errors; Operator Action Tree/Time Reliability Correlation*, NUREG/CR-3010 (BNL-NUREG-51601), U.S. Nuclear Regulatory Commission, November 1982.

Hannaman, G.W., Spurgin, A.J., Joksimovich, V., Wreathall, J., and Orvis, D.D., *Systematic Human Action Reliability Procedure (SHARP)*, NUS Corporation, San Diego. Prepared for the Electric Power Research Institute, Report NP-3583, 1984.

Hasegawa, H.K., Lambert, H.E., and Naanep, G.P., "Fire-Protection Study of the 2XIIB Mirror-Fusion Facility." Presented at the Fourth International System Safety Conference, July 9-13, 1979, San Francisco, California.

Heubach, W.F. and Philipson, L.L., *Investigation of LARA Failure Rate Methodology.* Prepared for U.S. Air Force Western Space and Missile Center, Technical Report No. 85-3135, J.H. Wiggins Co., September 1985.

Heubach, W.F. and Philipson, L.L., *Investigation of Approximations and Assumptions in New LARA Failure Rate Methodology*. Prepared for U.S. Air Force Western Space and Missile Center, Technical Report No. 86-3147-02, NTS Engineering, September 1986.

Higson, D.J., "Nuclear Reactor Safety Goals and Assessment Principles," *Nuclear Safety*, pp. 1-13, January-February 26, 1985.

Holloway, C.A., *Decision Making Under Uncertainty: Models and Choices*, Prentice Hall, Englewood Cliffs, N.J., 1979.

Huebel, J.G. and Myers, G.K., *Tables of Confidence Bounds for Failure Probabilities,* Lawrence Livermore Laboratory, UCRL-51990, January 1976.

Jeffreys, H., *Theory of Probability*, 3rd ed., Clarendon Press, Oxford, England, 1961.

Kahneman, D. and Tversky, A., "Prospect Theory: An Analysis of Decision Under Risk," *Econometrica*, Vol. 47, No. 2, March 1979.

Kaplan, S., "On a Two-Stage Bayesian Procedure for Determining Failure Rates from Experiential Data," *IEEE Transactions on Power Apparatus and Systems*, 1981a.

Kaplan, S., "On the Method of Discrete Probability Distributions in Risk and Reliability Calculations--Applications to Seismic Risk Assessment," *Risk Analysis,* Vol. 1, No. 3, 1981b.

Keeney, R.L., "Equity and Public Risk," *Operations Research*, Vol. 28, No. 3, Part I, May-June 1980.

Kelly, A.P. and Stillwell, D.W., *Application and Comparison of the GO Methodology and Fault Tree Analysis*, Pickard, Lowe and Garrick, Inc., Report PLG-0217, 1981.

Lambert, H.E., *Fault Trees for Decision Making in Systems Analysis*, Lawrence Livermore National Laboratory, UCRL 51829, 1975.

Lapp, S.A. and Powers, G.J., "Computer Aided Synthesis of Fault Trees," *IEEE Transactions on Reliability,* Vol. R-26, pp. 2-13, 1977a.

Lapp, S.A. and Powers, G.J., "The Synthesis of Fault Trees," in *Nuclear Systems Reliability Engineering and Risk Assessment,* Fussell, J.B. and Burdick, G.R. (eds.), Society for Industrial and Applied Mathematics (SIAM), 1977b.

Lewis, H.W. et al., *Risk Assessment Review Group Report to the U.S. Nuclear Regulatory Commission*, NUREG-CR-0400, September 1978.

Lindley, D.V. and Singpurwalla, N.D., "Reliability (and Fault Tree) Analysis Using Expert Opinions," *J. Amer. Stat. Assn.*, Vol. 81, pp. 87-90, March 1986.

Linnerooth, J., *The Evaluation of Life-Saving: A Survey*, International Institute for Applied Systems Analysis, Research Report RR-75-21, July 1975.

Lowrance, W.W., *Of Acceptable Risk - Science and the Determination of Safety*, William Kaufman, Inc., Los Altos, California, 1976.

Mann, N.R., "Simplified Expressions for Obtaining Approximately Optimum Series System-Reliability Confidence Bounds from Exponential Subsystem Data," *J. Amer. Stat. Assn.*, Vol. 69, pp. 492-495, June 1974.

Mann, N.R., Schafer, R.E., and Singpurwalla, *N.D., Methods for Statistical Analysis of Reliability & Life Data*, John Wiley & Sons, New York, 1974.

Marshall, A.W. and Olkin, I., "A Multivariate Exponential Distribution," *J. Amer. Stat. Assn.*, Vol. 62, pp. 30-44, 1967.

Martz, H.F. and Waller, R.A., *Bayesian Reliability Analysis*, John Wiley & Sons, New York, 1982.

Martz, H.F., and M. Bryson, "On Combining Data for Estimating the Frequency of Low-Probability Events with Application to Sodium Valve Failure Rates," *Nuclear Science and Engineering,* 1982.

Martz, H.F., et al., *Comparison of Methods for Uncertainty Analysis of Nuclear Power Plant Safety System Fault Tree Models*, NUREG/CR-3262, Los Alamos National Laboratory Report, 1983.

Maximus, Inc., *Handbook for the Calculation of Lower Statistical Confidence Bounds on System Reliability.* Prepared for U.S Army Research and Development Command, June 1981.

Morgan, J.M. and Andrews, J.D., *Assessment of Safety Systems Using Fault Tree Analysis, Midlands Research Station, British Gas Corporation, Communication 1242.* Presented at the 50th Autumn Meeting, Eastbourne, November 1984.

Mosleh, A., and Apostolakis, G., "Some Properties of Distributions Useful in the Study of Rare Events," *IEEE Transactions on Reliability*, 1982.

National Aeronautics and Space Administration, Headquarters Safety Division, *Safety Risk Management Program Plan*, Vols. I and II, April and June 1987.

Nielsen, S., *Use of Cause-Consequence Charts in Practical Systems Analysis,* Report Riso-M-1743, Danish AEC, September 1974 (also in *Reliability and Fault Tree Analysis*, Barlow et al. (eds.), SIAM, 1975).

Nuclear Regulatory Commission, *Reactor Safety Study: An Assessment of Accident Risks in U.S. Commercial Nuclear Power Plants*, U.S. Nuclear Regulatory Commission, WASH-1400 (NUREG-75-014), 1975.

Nuclear Regulatory Commission, *Guidelines for Control Room Design Reviews*, NUREG-0700, 1981.

Nuclear Regulatory Commission, *PRA Procedures Guide*, NUREG/CR-2300, January 1983.

Okrent D. and Whipple, C., *An Approach to Societal Risk Acceptance Criteria and Risk Management,* UCLA-ENG-7746, June 1977.

Otway, H.J. and Cohen, J.J., *Revealed Preferences: Comments on the Starr Benefit-Risk Relationships,* International Institute for Applied Systems Analysis, Research Memorandum RM-75-5, March 1975.

Parry, G.W. and Winter, P.W., "The Characterization and Evaluation of Uncertainty in Probabilistic Risk Analysis." *Nuclear Safety*, Vol. 12, January-February 1981.

Philipson, L.L. and Gasca, J.D., *Risk Assessment Methodologies and Their Uncertainties*, Vol. I, *A Review of Risk Estimation Approaches*, and Vol. II, *A Review of Risk Evaluation Approaches.* Report prepared for the National Science Foundation, J.H. Wiggins Co., March 1982.

Philipson, L.L. and Tran, D., *Investigation of LARA Failure Rate Methodology.* Prepared for U.S. Air Force Western Space and Missile Center, Technical Report No. 84-3123, J.H. Wiggins Co., September 1984.

Pratt, J.W., "A Normal Approximation for Binomial, F, Beta, and Other Common, Related Tail Probabilities, I (with D.B. Peizer) and II," *J. Amer. Stat. Assn.*, Vol. 63, pp. 1416-1483, 1968.

Rasmuson, M., Burdick, R., and Wilson, R., *Common Cause Analysis Techniques: A Review And Comparative Evaluation*, EG & G Idaho, Inc./USDOE, Tree-1349, September 1979.

Schwing, R.C., "Longevity Benefits and Costs of Reducing Various Risks," *Technological Forecasting and Social Change*, Vol. 13, pp. 333-345, 1979.

Sirvanci, M. and Yang, G., "Estimation of the Weibull Parameters Under Type I Censoring," *J. Amer. Stat. Assn.*, Vol. 79, pp. 183-187, 1984.

Slovic, P. and Fischoff, B., "How Safe is Safe Enough? Determinants of Perceived and Acceptable Risk." Paper of Decision Research, Eugene, Oregon, January 1979.

Spetzler, C.S., and Stael von Holstein, C.A.S, "Probability Encoding in Decision Analysis," *Management Science*, Vol. 22, pp. 340-358, 1975.

Stael von Holstein, C.A.S., *Assessment and Evaluation of Subjective Probability Distributions*, The Economic Research Institute, Stockholm School Of Economics, Stockholm, Sweden, 1970.

Starr, C., "Social Benefit Versus Technological Risk," *Science*, Vol. 165, pp. 1232-1238, September 1969.

Swain, A.D. and Guttman, H.E., *Handbook of Human Reliability Analysis With Emphasis on Nuclear Power Plant Applications*, Sandia National Laboratories, NUREG/CR-1278, 1983.

Tversky, A., and Kahneman, D., "Judgment Under Uncertainty: Heuristics and Biases," *Science*, Vol. 185, pp. 1124-1131, 1974.

U.S. Department of Defense, Military Standard, *Human Engineering Design Criteria for Military Systems, Equipments, and Facilities*, MIL-STD-1472C, 1981.

U.S. Department of Defense, Military Standard, *System Safety Program Requirements*, MIL-STD-882B, March 1986.

Vesely, W.E., "Estimating Common Cause Failure Probability in Reliability and Risk Analyses: Marshall-Olkin Specializations," *Proceedings, International Conference on Nuclear Systems Reliability Engineering and Risk Assessment,* Gatlinburg, Tennessee, June 1977.

Vesely, W.E., Roberts, N.H., Haasl, D.F., and Goldberg, F.F., *Fault Tree Handbook*, Nuclear Regulatory Commission, NUREG-0492, 1981.

Williams, L.R. and Gately, W.V., *GO Methodology: An Overview*, Electric Power Research Institute, Report NP-765, May 1978.

Winkler, R.L., "The Consensus of Subjective Probability Distributions," *Management Science*, Vol. 15, pp. B61-B75, 1968.

Winkler, R.L. and Hays, W.L., *Statistics: Probability, Inference, and Decision*, Second Edition, Holt, Reinhart and Winston, New York, 1975.

Winterbottom, A., "Lower Confidence Limits for Series System Reliability from Binomial Subsystem Data," *J. Amer. Stat. Assn.*, Vol. 69, pp. 782-788, September 1974.

Wreathall, J., "Current Development in Human Reliability Modeling," Informal Communication, Science Applications International Corporation, 1987.

Wright, F.T., Engelhardt, M., and Bain, L.J., "Inferences for the Two-Parameter Exponential Distribution Under Type I Censored Sampling," *J. Amer. Stat. Assn.*, Vol. 73, pp 650-655, 1978.

GLOSSARY

ACCEPTABLE RISK - A level of risk from a hazardous activity determined on the basis of specific criteria to be sufficiently low to enable the activity to be instituted or continued.

ACCIDENT RISK - A measure of the potential for loss, damage, or injury from mishaps in a hazardous activity or element thereof.

APPROPRIATE (RISK ASSESSMENT PROCEDURE) - A procedure tailored to the hazards evaluation and disposition decision-making needs of an activity, taking into account the significance of the risks of the activity, the resources required to carry out their assessment, and the capability for making available the data required by the assessment.

AVAILABILITY - A measure of the likelihood that an item will operate when called upon.

BASIC (SOMETIMES, PRIMARY) EVENT - The occurrence of a fault or failure in a system component, or the occurrence of an external event, that can initiate, or participate in, a sequence of events leading to a mishap.

BAYESIAN INFERENCE OR ESTIMATION - The combination in accordance with Bayes' Theorem of a "prior" model based on related information (engineering judgment, similar test or observational data, etc.) for the probability distribution of the possible values of a parameter, with such specific statistical data as are available, to generate a "posterior" model for the parameter's distribution. It is applied to develop parameter value inferences when specific data are too sparse for classical statistical methods or otherwise in order to make optimal use of all information relevant to the parameter.

CATASTROPHIC CONDITION - A hazardous condition that may cause death, major injuries, or major damage.

CATASTROPHIC FAILURE - A failure that may result in a catastrophic condition.

CONSEQUENCE - A possible harmful outcome of a failure or of a mishap, depending on the context.

CREDIBLE CONDITION - A condition that has a reasonable likelihood of occurrence.

CRITICAL CONDITION - A hazardous condition that may lead to a mishap with severe injuries or illnesses, or major damage to property or equipment.

EFFECT - A result of an undesired event such as a mishap; for example, the release and dispersion of a given quantity of hazardous energy or material.

ERROR - A discrepancy between a computed, observed, or measured value or condition and the true, specified, or theoretically correct value or condition.

EXPOSURE - Depending on the context: (a) population, property, or other value system exposed to possible harm by a given hazardous activity; (b) the operating or test time, cycles, or other measure of the opportunity for failure or mishap events to occur.

FAILURE - The inability of a system, subsystem, component or part to perform its required function within specified limits, under specified conditions, for a specified duration.

FAILURE CAUSE - The physical or chemical process, design defect, quality defect, part misapplication, or other factors or events which are the basic reasons for a failure or which initiate the physical process by which deterioration proceeds to a failure.

FAILURE EFFECT - The consequence of a failure on the operation, function, or status of a component, assembly, or system.

FAULURE MODE - A particular way in which a failure can occur.

FAILURE MODES AND EFFECTS ANALYSIS - A systematic "bottom-up" analysis performed to identify and document all identifiable failure modes at a prescribed level of assembly, and to specify the resultant effects of the failure modes at higher levels.

FAILURE RATE - The number of failures per unit time or other measure of the opportunity for failures to occur.

FAULT TREE - A "top-down" logic modeling technique, in which an undesired system-level event is specified and the system is then analyzed in the context of its operational environment to identify all credible ways in which the undesired event can occur.

FAULT TREE ANALYSIS - A logical analysis of all events and their interrelationships that can cause an undesired system-level event.

HAZARD - An existing or potential condition that gives rise to a risk or risks of harm.

HAZARD ANALYSIS - The identification and evaluation of existing and potential sources of risks of harm and the recommendation of mitigations for the sources found.

HAZARD SEVERITY - An assessment of the worst-case harm from credible mishaps that could occur because of a given hazard.

HAZARDOUS EVENT - An event whose occurrence has potential for harm.

HAZARDOUS FACILITY - A NASA or NASA contractor facility dedicated to research and technology development whose activities involve hazards with significant or potentially significant risks of harm to people or property.

HAZARDOUS OPERATION - An operation involving activities with risks of injury or loss of life to personnel, of damage to systems/equipment, or of harmful environmental impacts.

LOGIC (OR LOGIC TREE) MODEL - One of several forms of the graphical exposition of the functional states of system elements and their interactions which contribute to the occurrence of a given system-level event. The model is a reliability-oriented model if the event is the performance of a desired function; it is a failure-oriented model if the event is the occurrence of a system failure or mishap.

LOSS - An outcome of a mishap, expressed in terms such as number of injuries or cost of damage.

MEAN-TIME-BETWEEN (OR TO)-FAILURES - The statistical mean of the distribution of times between (or to) failures.

MEAN-TIME-TO-RESTORE (OR REPAIR) - The statistical mean of the distribution of times-to-restore (or repair).

MISHAP - An unplanned event that results in death, injury, illness, or damage to equipment, property, or the environment, or in a mission or test failure that has significant program impact or visibility. Mishap is synonymous with NASA mishap and NASA contractor mishap.

MISSION FAILURE - A NASA mishap that prevents accomplishment of a primary mission objective.

OPERATING AND SUPPORT HAZARD ANALYSIS - An analysis of procedurally controlled activities during all phases of operation of a system or facility performed to identify hazards and recommend risk reduction alternatives.

PRELIMINARY HAZARD ANALYSIS - An initial identification and evaluation of hazards in a system concept or preliminary design.

PROBABILITY - The likelihood that a given event will occur per unit opportunity (time, cycles, etc.) for its occurrence.

REDUNDANCY - The availability of more than one means to accomplish a given function where more than one means must fail before the function fails.

RELIABILITY - A characteristic of a component or system, expressed as a probability that it will perform its required functions under defined conditions at designated times for specified operating periods.

RISK (SAFETY) - The chance of occurrence of a loss or mission failure in the operation of a system or facility. It is a function of the possible frequency of occurrence of a mishap, of the potential severity of the mishap's consequences, and of the uncertainties associated with the frequency and severity.

RISK ACCEPTANCE - The acceptance by an individual or organization of the level of risk that has been assessed as accruing to a given activity.

RISK ASSESSMENT - The process of qualitative risk categorization or quantitative risk estimation, followed by the evaluation of risk significance.

RISK CONTRIBUTORS LIST - A listing of identified hazards in an activity in accordance with the magnitudes of the risks of harm to which they give rise.

RISK DISPOSITION - The decision or the results thereof on the treatment of an identified risk (accept, tolerate through waivers, or mitigate by specified means).

RISK MANAGEMENT - The process of balancing risk with cost, schedule, and other programmatic considerations. It consists of risk identification, risk assessment, decision-making on the disposition of risk, and tracking the effectiveness of the results of the action resulting from the decision.

RISK MANAGEMENT ASSURANCE AND SUPPORT - The oversight of, and assistance to, risk management activities to ensure their effective conduct.

RISK MITIGATION - A means for eliminating or modifying a hazard so as to diminish the risks of harm to which it gives rise.

SIGNIFICANT OR POTENTIALLY SIGNIFICANT RISK - A risk whose likelihood and/or severity components, considered together, cannot immediately be decided to be acceptable by cognizant authority.

SYSTEM - An integrated assembly of equipment, procedures, and skills that performs or supports an operational mission.

SYSTEM (SUBSYSTEM) HAZARD ANALYSIS - A system (subsystem)-level hazard analysis that provides a comprehensive evaluation of the risk assumed when each subsystem (assembly) is put into operation.

SYSTEM SAFETY - A principal activity in support of safety risk management in which engineering and management principles, criteria, and techniques are applied to maximize safety within given constraints of operational effectiveness, time, and cost throughout all phases of a system's life cycle.

SYSTEM SAFETY ANALYSIS - Systematic mishap risk and risk mitigations identification and evaluation through a series of iterative qualitative or quantitative analyses.

TOLERABLE RISK - A risk decided by cognizant authorities to be adequately in balance with the benefits of the activity giving rise to the risk. A tolerable risk may not be a formally acceptable risk.

UNCERTAINTY - The extent to which a true state such as the value of a parameter can vary from its assessment or prediction. For a parameter value, if this possible variation is quantified as a range or interval, uncertainty is defined by upper and lower limits (in classical statistics, deriving from sample data, confidence limits) about an estimated value within which the true value falls with a given probability.

WAIVER - The documented decision by cognizant authority to tolerate the (safety) risk in using or accepting an article that does not meet specified requirements.