Fault Trees for Diagnosis of System Fault Conditions

Howard E. Lambert

Lawrence Livermore Laboratory, P. O. Box 808, Livermore, California 94550

and

George Yadigaroglu

University of California, Department of Nuclear Engineering, Berkeley, California 94720

Received May 10, 1976
Revised August 2, 1976

Methods for generating repair checklists on the basis of fault tree logic and probabilistic importance are presented. A one-step-ahead optimization procedure, based on the concept of component criticality, minimizing the expected time to diagnose system failure is outlined. Options available to the operator of a nuclear power plant when system fault conditions occur are addressed. A low-pressure emergency core cooling injection system, a standby safeguard system of a pressurized water reactor power plant, is chosen as an example illustrating the methods presented.

I. INTRODUCTION

Fault tree analysis (FTA) is a method of system safety analysis that evolved from the aerospace industry in the early 1960's. A fault tree is a Boolean-logic model that depicts the parallel and sequential combinations of events that cause a top event to occur. The top event usually represents an undesired system state or hazardous condition. The first step in FTA is fault tree construction, i.e., the systematic identification of links and relationships between system components that could produce the conditions necessary for the occurrence of the top event and graphical display according to a standard format. Once constructed, the fault tree can be evaluated either qualitatively or quantitatively, depending on the scope, extensiveness, and use of the analysis. An important step in a qualitative evaluation is the determination of the minimal cut sets, i.e., the set of basic events whose occurrence causes the top event to occur. The basic events are events such as generic hardware failures, human error, or environmental conditions, and they represent the limit of resolution in the fault tree. Quantitative evaluations are concerned with the determination of probabilistic characteristics of the fault tree such as the probability of the top event or the probabilistic importance of basic events or minimal cut sets. The reader should consult Refs. 1 through 9 for a general discussion of FTA and

References 10 through 14 for the probabilistic evaluation of fault trees.

Traditionally, FTA has been used as a design tool. It can identify potential accidents in system design and can help eliminate costly design changes and retrofits. More recently, it has been used as a tool for evaluating the safety of existing complex systems such as nuclear power plants. In this paper, FTA is used as a diagnostic tool. We show how FTA can predict the most likely causes of failure in the event of a system breakdown. The prediction or diagnosis is accomplished with maximum economy of time and effort, since learning by experience is included in the analytical framework.

To accomplish these objectives, the following topics are covered:

1. concepts of component criticality and probabilistic importance
2. generation of repair checklists on the basis of fault tree logic
3. system diagnosis under a time constraint.

As an example, we choose the low-pressure emergency coolant injection system (LPIS) of a pressurized water reactor (PWR) nuclear power plant for which we generate repair checklists.

We conclude the paper by considering the options available to an operator when system fault conditions occur and how the best future course of plant operation can be determined on the basis of a risk assessment. In particular, we consider the decision regarding shutdown of a PWR plant when the LPIS is found partly inoperable during plant operation.

II. PROBABILITY IMPORTANCE

Some probabilistic aspects of FTA are discussed in this section to introduce the concepts of probabilistic importance. By computing probabilistic importance, we generate a numerical ranking of events in a fault tree. Then, we can determine the component malfunctions and fault conditions most likely to be contributing to the occurrence of the specified top event. As we examine the system, we can incorporate the knowledge we gain into the importance rankings, making FTA a truly interactive diagnostic tool.

The purpose of this section is to:

1. present the mathematical notation used in the remainder of the paper
2. explain the concept of component criticality
3. briefly summarize the concepts of probabilistic importance.

The reader should consult Refs. 13, 15, 16, and 17 for a more detailed treatment of the subject of probabilistic importance.

II.A. Mathematical Notation (Ref. 18)

Probabilistic definitions of importance can be expressed in terms of a g function that establishes the probability of the top event in terms of basic-event probabilities. To generate this function, we need a Boolean expression for the top event in terms of the Boolean variables of the basic events. We describe the occurrence or nonoccurrence of each basic event at time t by an indicator variable, \( Y_i(t) \)

\[
Y_i(t) = \begin{cases} 
1 & \text{if basic event } i \text{ has occurred at time } t \\
0 & \text{otherwise} 
\end{cases}
\]

The top event is assigned an indicator variable or structure function, \( \psi(Y(t)) \), such that

\[
\psi(Y(t)) = \begin{cases} 
1 & \text{if the top event has occurred at time } t \\
0 & \text{otherwise} 
\end{cases}
\]

where

\[
Y(t) = [Y_1(t), Y_2(t), \ldots, Y_n(t)]
\]

is the vector of basic-event indicator variables at time \( t \) and \( n \) is the number of basic events in the fault tree.

There are two operators that can be used to express \( \psi(Y(t)) \) in terms of \( Y(t) \): the AND operator, \( \Pi \), and the OR operator, \( \Pi \). The AND structure function is given by

\[
\psi(Y(t)) = \prod_{i=1}^{n} Y_i(t) = Y_1(t) \cdot Y_2(t) \cdot \ldots \cdot Y_n(t)
\]
while the structure function for an OR gate with \( n \) inputs is
\[
\psi[Y(t)] = \prod_{i=1}^{n} Y_i(t) \equiv 1 - \prod_{i=1}^{n} [1 - Y_i(t)].
\]

At this point it is convenient to define basic events in the fault tree as component failures and to represent the top event as a system failure. With this convention, \( Y_i = 1 \) denotes failure of component \( i \). The structure function for a parallel system is an AND structure function since all components must fail to cause the system to fail. For a series system, an OR structure function must be used.

Component \( i \) is critical to system failure at time \( t \) if the system is in a state denoted as \( Y(t) \) such that the system makes a transition from the unfailed state to a failed state when component \( i \) fails. In terms of the notation previously described, the above statement implies
\[
\psi[Y_1(t), Y(t)] - \psi[0, Y(t)] = \begin{cases} 1 & \text{if component } i \text{ is critical to system failure at time } t \\ 0 & \text{otherwise} \end{cases},
\]
where the meaning of the abbreviated notation used is
\[
egin{align*}
[1_i, Y(t)] &= [Y_1(t), \ldots, Y_{i-1}(t), 1, Y_{i+1}(t), \ldots, Y_n(t)] \\
[0_i, Y(t)] &= [Y_1(t), \ldots, Y_{i-1}(t), 0, Y_{i+1}(t), \ldots, Y_n(t)].
\end{align*}
\]

In a parallel system, for a component to be critical to system failure at time \( t \), we must have
\[
\psi[Y_1(t), Y(t)] - \psi[0, Y(t)] = \prod_{j=1 \atop j \neq i}^{n} Y_j(t) = 1,
\]
which implies that \( Y_j(t) = 1 \) for all \( j \neq i \). Thus, in a parallel system for a component to be critical to system failure at time \( t \), all the remaining components must have failed at time \( t \).

For a series system, we must have
\[
\psi[Y_1(t), Y(t)] - \psi[0, Y(t)] = \prod_{j=1 \atop j \neq i}^{n} [1 - Y_j(t)] = 1
\]
if \( Y_j(t) = 0 \) for all \( j \neq i \).

Note that for statistically independent basic events, the \( g \) function is obtained by eliminating powers of indicator variables and then merely
\[\begin{align*}E[\ Y_i(t)\] & = \text{Prob}[Y_i(t) = 1] \equiv q_i(t)\ .
\]

If the basic event, \( i \), describes failure of an unrepairable component, then \( q_i(t) \) is given by
\[q_i(t) = 1 - \exp \left[ - \int_0^t \lambda_i(t') dt' \right],\]
where \( \lambda_i(t')dt' \) is the conditional probability of failure in \((t', t' + dt')\) and is commonly referred to as the failure rate. If repair of the failed component is permitted, either by scheduled or off-schedule maintenance, then the probability that component \( i \) is failed at time \( t \) is given by its unavailability, \( \overline{A}_i(t) \),
\[q_i(t) = \frac{\tau_i}{\tau_i + \mu_i},\]
which represents the fraction of time component \( i \) is expected to be failed at time \( t \). For off-scheduled maintenance, the asymptotic value of the unavailability is given by
\[\lim_{t \to \infty} \overline{A}_i(t) = \frac{\tau_i}{\tau_i + \mu_i},\]
where \( \tau_i \) is the mean time to repair for component \( i \) and \( \mu_i \) is the mean time to failure for component \( i \).

We now consider the top event or system characteristics. If the basic events are statistically independent, then the probability that the top event occurs by time \( t \) is given by the \( g \) function mentioned previously, \( g[q(t)] \). It can be recognized that
\[E[\psi[Y(t)]] = \text{Prob}[\psi[Y(t)] = 1] = g[q(t)].\]

The system reliability is simply \( 1 - g[q(t)] \). For example, for a parallel system of two components, \( g[q(t)] \) is given by
\[g[q(t)] = q_1(t) \cdot q_2(t),\]
and the probability that component 1 is critical to system failure is
\[g[1, q_2(t)] = g[0, q_2(t)] = q_2(t).\]
For a series system of two components,
\[g[q(t)] = 1 - [1 - q_1(t)][1 - q_2(t)]\]
and the probability that component 1 is critical to system failure is
\[g[1, q_2(t)] = g[0, q_2(t)] = 1 - q_2(t).\]

Note that for statistically independent basic events, the \( g \) function is obtained by eliminating powers of indicator variables and then merely

---

19In off-schedule maintenance, repair (or replacement) of the failed component takes place immediately on detection of the failure, whereas in scheduled maintenance, repair occurs at the end of some inspection interval.
substituting \( q_i(t) \) for \( Y_i(t) \) in the Boolean structure function.

The above results are well established in reliability theory (see, for example, Refs. 10, 11, and 20).

II.B. Probabilistic Importance

We now consider probabilistic expressions for ranking basic events according to their importance in a fault tree. For the remainder of the paper, we assume that all basic events are statistically independent. Further, we assume that the system structure function is coherent, i.e., \( \psi(Y(t)) \) is a monotonic Boolean function. The implications of this statement are discussed in Ref. 11.

We introduce two measures of importance computed in terms of \( g_i(q(t)) \), a function that describes the state of the system at one point in time.

In 1969, Birnbaum introduced the concept of importance for coherent systems. He defined the reliability importance of a component \( i \) as the rate at which system reliability improves as the reliability of component \( i \) improves. If by following our convention we construct a fault tree where the top event is system failure and the basic events are component failures, then the Birnbaum definition of component importance becomes

\[
\frac{\partial g_i(q(t))}{\partial q_i(t)} = g_i(1, q(t)) - g_i(0, q(t)) \quad \text{def} \quad \Delta g_i(t) .
\]

The partial derivative relationship holds by the idempotency law for Boolean variables, i.e., \( \psi(Y(t)) \) is a linear function of \( Y_i(t) \), which implies that \( g_i(q(t)) \) is linear in \( q_i(t) \) when independence is assumed. Stated in other terms, the Birnbaum measure of importance is the probability of a component being critical to system failure. It is possible that by the time system failure is observed, more than one min cut sets could have failed. In this case, restoring a failed component to a working state does not necessarily restore the system to a working state. In other words, it is possible that failure of a component can be contributing to system failure without being critical. (Component \( i \) is contributing to system failure if a min cut set containing \( i \) has failed.) If we let

\[
\psi_i^C(Y(t)) = \text{Boolean structure function for the union of all cut sets that contain basic event, } i
\]

then the probability that component \( i \) is contributing to system failure is given by

\[
E[\psi_i^C(Y(t))] \defeq g_i'(q(t)) ,
\]

and

\[
I_{FV} = \frac{g_i(q(t))}{g(q(t))} \quad \text{def} \quad I_{FV}(t) .
\]

This concept of importance was first proposed by Vesely and Fussell who introduced it in the literature. Both the Birnbaum and Fussell-Vesely measures of importance are used here.

As an example of the Fussell-Vesely measure of importance, we consider a series and a parallel system of two components. In a series system, there are two min cut sets, (component 1 fails) and (component 2 fails), simply denoted as (1) and (2). The probability that component 1 is contributing to system failure given that system failure has occurred is given by

\[
I_{FV} = \frac{q_1(t)}{q_1(t)q_2(t)} = 1 .
\]

Other measures of importance can be defined to describe system failure in terms of sequences of component failures. These are discussed by Barlow and Proschan and by Lambert.

III. GENERATION OF REPAIR CHECKLISTS

In this section we present methods by which repair checklists can be generated. If a fault tree can accurately simulate system failure, i.e., if all failures can be described in terms of Boolean logic, then the fault tree can be quantitatively evaluated to determine the critical events. In the event of system or subsystem breakdown, a repair checklist can be generated for an operator to follow in diagnosing the particular system failure. The basic events on this checklist can be ordered according to their importance when system failure occurs.

III.A. Checklist Generation Scheme

The order in which the components are listed on the checklist should reflect the knowledge that the operator gains about the system as he sequentially examines each component in the checklist. The ranking of the basic events should be done on a conditional basis, i.e., in the framework of

\[22\text{W. E. VESELY, Division of Reactor Safety Research, U.S. Nuclear Regulatory Commission, Private Communication (1972).}\]
learning by experience. For example, if the operator finds that the first event on the checklist has not occurred, then the second event on the checklist should be the most critical to system failure, given that the first event has not occurred. In general, the $i$th event is most critical to system failure, given that the first ($i - 1$) events have not occurred.

If a component, say $i$, in the checklist is found to be failed and is contained in a cut set of order two or higher, then a sublist is generated for component $i$. In this sublist, we rank the cut sets containing the failed component, $i$, according to their importance computed on a conditional basis. The operator then checks the components in the cut sets that contain component $i$.

In general, it is unwise to include the triple or higher order cut sets in the sublist since for maintained or inspected systems, the simultaneous occurrence of three independent events is rare. It is felt that the criteria adopted by the Reactor Safety Study\(^7\) also are valid for checklist generation. According to these criteria, we retain the most important cut sets only:

1. single passive faults
2. single active faults
3. double active faults.\(^7\)

If these criteria are adopted, the sublist becomes a list of active components ranked according to their probability of occurrence. The advantage of keeping only the most important cut sets is that a multitude of trivial combinations that are generated normally from typical fault trees are eliminated from consideration. It must be kept in mind that the purpose of the checklist is to aid the operator in making decisions under a time constraint.

Although all basic events are assumed to be independent, dependent failures can be incorporated into the scheme outlined above. This is accomplished by including basic events that cause secondary failures, i.e., environmental or operational conditions capable of simultaneously failing two or more system components. When we check for these secondary failure conditions, we generate a sublist of the components sensitive to these conditions.

In generating the checklist, the possibility of false alarms should also be taken into account by considering the reliability of the monitoring device that indicated system failure.

**III.B. Checklist Generation for the LPIS**

The system that is chosen for checklist generation is the LPIS of a PWR power plant. The LPIS is a standby safety system, part of the emergency core cooling system (ECCS). A simplified piping schematic of this system is shown in Fig. 1. Following a loss-of-coolant accident (LOCA), the LPIS operation is considered successful when at least one of its legs discharges water continuously at a rate of 3000 g/min and a pressure of 300 psi into the cold legs of the reactor.

Part of the control circuit that actuates the LPIS is shown in Fig. 2, which includes a brief system description. For simplicity, it is assumed that this system is tested by closing a switch (not shown) that energizes relays that in turn close contacts P1, P2, and P3. A checklist is generated for leg A for the case when the pressure gauge fails to indicate 300 psi when the test switch is closed. This checklist contains the list of events that are most likely to have occurred under those conditions.

Part of the fault tree\(^{13}\) that illustrates the failure of leg A of the LPIS system is shown in Fig. 3. The remainder of the fault tree consists of four additional pages and is omitted for sake of space. However, a Boolean equivalent of the entire fault tree is shown in Fig. 4. In Fig. 4, the numbers that are listed as gate inputs represent basic events that are listed alphabetically in Table I. The basic events in the LPIS fault tree are coded according to a seven-digit alphanumeric designator following the practice adopted for the Reactor Safety Study.\(^7\) The first digit indicates the event type, X represents a circle, and Z represents a diamond. The second and third digits indicate the component type, e.g., PM stands for pump, MV for motor-operated valve. The fifth and sixth digits identify the specific events or components listed in the event description of Table I. The seventh digit represents the failure mode of the component; e.g., Q stands for short-circuit, A stands for "does not start," etc. Note
# Fault Tree Diagnosis

## Table I

<table>
<thead>
<tr>
<th>Event</th>
<th>Number</th>
<th>Event Description</th>
<th>Unavailability</th>
</tr>
</thead>
<tbody>
<tr>
<td>XCB01K</td>
<td>1</td>
<td>Circuit breaker contacts fail to close</td>
<td>$10^{-3}/d$</td>
</tr>
<tr>
<td>XCV01C</td>
<td>2</td>
<td>Check valve A jammed closed</td>
<td>$10^{-4}/d$</td>
</tr>
<tr>
<td>XCV02C</td>
<td>3</td>
<td>Check valve B jammed closed</td>
<td>$10^{-4}/d$</td>
</tr>
<tr>
<td>XMV01D</td>
<td>4</td>
<td>MOV 1 fails to open</td>
<td>$10^{-4}/d$</td>
</tr>
<tr>
<td>XMV02D</td>
<td>5</td>
<td>MOV 2 fails to open</td>
<td>$10^{-5}/d$</td>
</tr>
<tr>
<td>XPD01K</td>
<td>6</td>
<td>Pressure transducer contacts P1 fail to close</td>
<td>$10^{-5}/d$</td>
</tr>
<tr>
<td>XPD02K</td>
<td>7</td>
<td>Pressure transducer contacts P2 fail to close</td>
<td>$10^{-5}/d$</td>
</tr>
<tr>
<td>XPD03K</td>
<td>8</td>
<td>Pressure transducer contacts P3 fail to close</td>
<td>$10^{-5}/d$</td>
</tr>
<tr>
<td>XPM01A</td>
<td>9</td>
<td>Pump motor fails to start</td>
<td>$10^{-5}/d$</td>
</tr>
<tr>
<td>XRE01K</td>
<td>10</td>
<td>#1 contacts fail to close</td>
<td>$10^{-4}/d$</td>
</tr>
<tr>
<td>XRE02K</td>
<td>11</td>
<td>#4 contacts fail to close</td>
<td>$10^{-4}/d$</td>
</tr>
<tr>
<td>XRE03K</td>
<td>12</td>
<td>#2 contacts fail to close</td>
<td>$10^{-4}/d$</td>
</tr>
<tr>
<td>XRE04K</td>
<td>13</td>
<td>#5 contacts fail to close</td>
<td>$10^{-4}/d$</td>
</tr>
<tr>
<td>XRE05K</td>
<td>14</td>
<td>#7 contacts fail to close</td>
<td>$10^{-4}/d$</td>
</tr>
<tr>
<td>XRE06K</td>
<td>15</td>
<td>#3 contacts fail to close</td>
<td>$10^{-4}/d$</td>
</tr>
<tr>
<td>XRE07K</td>
<td>16</td>
<td>#6 contacts fail to close</td>
<td>$10^{-4}/d$</td>
</tr>
<tr>
<td>XRE01Kb</td>
<td>17</td>
<td>Manual valve fails to open</td>
<td>$10^{-4}/d$</td>
</tr>
<tr>
<td>ZBS01N</td>
<td>18</td>
<td>No power on bus 480 1H</td>
<td>$(5 \times 10^{-4}) (A)$</td>
</tr>
<tr>
<td>ZBS02N</td>
<td>19</td>
<td>No power on bus DC1A</td>
<td>$(5 \times 10^{-5}) (A)$</td>
</tr>
<tr>
<td>ZBS03N</td>
<td>20</td>
<td>No power on bus MCC1H1-1</td>
<td>$(5 \times 10^{-5}) (A)$</td>
</tr>
<tr>
<td>ZCB02O</td>
<td>21</td>
<td>Circuit breaker #1 open</td>
<td>$7.2 \times 10^{-4}$</td>
</tr>
<tr>
<td>ZCB03O</td>
<td>22</td>
<td>Circuit breaker #2 open</td>
<td>$7.2 \times 10^{-4}$</td>
</tr>
<tr>
<td>ZMV03C</td>
<td>23</td>
<td>N.O. MOV 3 inadvertently closes</td>
<td>$7.2 \times 10^{-4}$</td>
</tr>
<tr>
<td>ZPF01P</td>
<td>24</td>
<td>Piping in leg A plugged</td>
<td>$7.2 \times 10^{-4}$</td>
</tr>
<tr>
<td>ZPF02C</td>
<td>25</td>
<td>Piping from RWST plugged</td>
<td>$7.2 \times 10^{-4}$</td>
</tr>
<tr>
<td>ZPF01R</td>
<td>26</td>
<td>Rupture in leg A of LPIS</td>
<td>$7.2 \times 10^{-4}$</td>
</tr>
<tr>
<td>ZPF02R</td>
<td>27</td>
<td>Rupture in pipe from RWST</td>
<td>$7.2 \times 10^{-4}$</td>
</tr>
<tr>
<td>ZTR01O</td>
<td>28</td>
<td>Open circuit or short circuit transfer #1</td>
<td>$7.2 \times 10^{-4}$</td>
</tr>
<tr>
<td>ZTR02O</td>
<td>29</td>
<td>Open circuit or short circuit transfer #2</td>
<td>$7.2 \times 10^{-4}$</td>
</tr>
<tr>
<td>ZWB01O</td>
<td>30</td>
<td>O.C. or S.C. in cable from LP1 PP to bus 4801H</td>
<td>$7.2 \times 10^{-4}$</td>
</tr>
<tr>
<td>ZWR02O</td>
<td>31</td>
<td>O.C. or S.C. in wiring of close coil circuit</td>
<td>$7.2 \times 10^{-4}$</td>
</tr>
<tr>
<td>ZWR03O</td>
<td>32</td>
<td>O.C. or S.C. in wiring of K1 coil circuit</td>
<td>$7.2 \times 10^{-4}$</td>
</tr>
<tr>
<td>ZWR04O</td>
<td>33</td>
<td>O.C. or S.C. in wiring of K4 coil circuit</td>
<td>$7.2 \times 10^{-4}$</td>
</tr>
<tr>
<td>ZWR05O</td>
<td>34</td>
<td>O.C. or S.C. in cable from MOV-1 to bus MCC1H1-1</td>
<td>$7.2 \times 10^{-4}$</td>
</tr>
<tr>
<td>ZWR06O</td>
<td>35</td>
<td>O.C. or S.C. in wiring of K2 coil circuit</td>
<td>$7.2 \times 10^{-4}$</td>
</tr>
<tr>
<td>ZWR07O</td>
<td>36</td>
<td>O.C. or S.C. in wiring of K5 coil circuit</td>
<td>$7.2 \times 10^{-4}$</td>
</tr>
<tr>
<td>ZWR08O</td>
<td>37</td>
<td>O.C. or S.C. in wiring of K7 coil circuit</td>
<td>$7.2 \times 10^{-4}$</td>
</tr>
<tr>
<td>ZWR09O</td>
<td>38</td>
<td>O.C. or S.C. in cable from MOV-2 to bus MCC1H1-1</td>
<td>$7.2 \times 10^{-4}$</td>
</tr>
<tr>
<td>ZWR10O</td>
<td>39</td>
<td>O.C. or S.C. in wiring of K2 coil circuit</td>
<td>$7.2 \times 10^{-4}$</td>
</tr>
<tr>
<td>ZWR11O</td>
<td>40</td>
<td>O.C. or S.C. in wiring of K5 coil circuit</td>
<td>$7.2 \times 10^{-4}$</td>
</tr>
<tr>
<td>ZXR01Y</td>
<td>41</td>
<td>Maintenance crew inadvertently closes manual valve</td>
<td>$10^{-4}/d$</td>
</tr>
</tbody>
</table>

$d = \text{demand}; \Delta t/T_m = \text{the fractional downtime due to maintenance.}$

A basic event whose occurrence can cause Leg A of the LPIS to fail on demand.

A basic event whose occurrence can cause the LPIS to fail on demand.

$A = \text{limiting unavailability.}$
that human error is also included in the fault tree as a cause of LPIS failure. For example, the basic event Z XV01Y represents the event "operator (or maintenance crew) inadvertently closes the manual valve."  

The unavailabilities of all the basic events in the fault tree of Fig. 3 are listed in Table I. The unavailability of all active components required to change state is given by their cyclic failure rate, i.e., the probability of failing to change state on demand. The emergency power buses are continuously operating systems; their unavailabilities are given by their limiting asymptotic values, $\bar{A}$. The unavailability of all other components is given by the product $\lambda \tau$, where $\lambda$ is an hourly failure rate and $\tau$ is the effective exposure time or fault duration time. The possibility of maintenance of these components is allowed since the entire LPIS is periodically tested. The events are listed by decreasing values if the Fussell-Vesely measure of importance in Table II. Inspection of this table shows that pump A failing to start has the greatest probability of causing failure of the LPIS to start. If pump A is working satisfactorily, we should check the circuit breaker for pump A. We see that events 23 and 30 are of equal importance and should be checked next. At this point in the checklist, we check for failure of a quasi-static component, i.e., a cable failure. At this point it is decided to check for a false alarm, since failure of an active component--in this case a pressure gauge--is more likely to occur than failure of a passive component. The components listed eighth in order in Table II are components contained in second-order cut sets. (Components in Table II with no asterisks are contained in second-order cut sets.) A sublist for motor-operated valve #1 is given in Table III. It is simply a listing of basic events contained in the same cut sets as MOV 1, listed in order of their probability of occurrence. For this particular example, the iteration process described previously was used to generate a checklist based on conditional probabilities. This checklist ranked components according to the same checking order as the initial listing of probabilistic importances in Table II. However,
Fig. 3. Fault tree for the LPIS.
by varying the values of the unavailabilities of components in Table I, for this example it was shown that the ranking done on a conditional basis differed from the ranking done on the initial listing. In general, the iteration process is necessary for ordering events on the checklist.

A useful feature of the checklist generation scheme is that a library of checklists can be prepared ahead of time so that a computer is not needed at the time the diagnosis is made. As time progresses, the checklists may have to be revised periodically to reflect increased knowledge concerning the system.

IV. SYSTEM DIAGNOSIS UNDER A TIME CONSTRAINT

In Sec. III, we generated repair checklists solely on the basis of probabilistic importance. We did not consider the time required to check components. In some cases, there may be a considerable risk or system degradation while a system or subsystem is down. In this section, we propose a checking scheme based on the concept of component criticality that minimizes the expected time required to diagnose system failure.

The scheme makes use of an expression that is a function of the component checking times as well as their probabilistic importance. Under the proposed scheme, if all components in the system had equal checking times, then the component that has the highest probability of being critical to system failure would be checked first. However, later in this section we show that the order in which the components are checked may change if component checking times are unequal.

It is assumed that the system is failed and that all causes of failure, i.e., basic events, in the fault tree are statistically independent. Under the proposed scheme, we check components one by one until failure of all the elements of at least one min cut set is confirmed. If after the cut set(s) is (are) repaired, system failure persists, we continue to check until another failure of a min cut set is observed. We continue in this manner until no more failed cut sets are found.

Each time we check a component in the system, there are three possible outcomes:

1. The component has not failed.
2. The component has failed but is not critical to system failure.
TABLE II
Importance Listing and Checklist for LPIS
(Probability that the LPIS fails on demand = 5.0 \times 10^{-4}; probability that the LPIS fails on demand when pump A has failed = 5.0 \times 10^{-3}).

<table>
<thead>
<tr>
<th>Order</th>
<th>Event Number</th>
<th>Event</th>
<th>Importance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9</td>
<td>XPM01A(^a)</td>
<td>4.380E-01(^b)</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>XCB01Ka</td>
<td>1.248E-01</td>
</tr>
<tr>
<td>3</td>
<td>23</td>
<td>ZMV03C</td>
<td>8.948E-02</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Check for false alarm</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>18</td>
<td>ZBS01Na</td>
<td>6.238E-02</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>XCV01Ca</td>
<td>1.247E-02</td>
</tr>
<tr>
<td>6</td>
<td>26</td>
<td>ZPP01Ra</td>
<td>8.978E-03</td>
</tr>
<tr>
<td>7</td>
<td>24</td>
<td>ZPP01Pa</td>
<td>5.487E-03</td>
</tr>
<tr>
<td>8</td>
<td>5</td>
<td>XMOV02D</td>
<td>3.166E-03</td>
</tr>
</tbody>
</table>

\(^a\) A basic event whose occurrence can cause leg A of the LPIS to fail on demand.
\(^b\) Read as 4.380 \times 10^{-1}.
\(^c\) A basic event whose occurrence can cause the LPIS to fail on demand.

TABLE III
Sublist for Motor Operated Valve #1

<table>
<thead>
<tr>
<th>Order</th>
<th>Event Number</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5</td>
<td>XMOV02D</td>
</tr>
<tr>
<td>2</td>
<td>22</td>
<td>ZCB03O</td>
</tr>
<tr>
<td>2</td>
<td>29</td>
<td>ZTR02O</td>
</tr>
<tr>
<td>3</td>
<td>38</td>
<td>ZWR090</td>
</tr>
<tr>
<td>3</td>
<td>15</td>
<td>XRE06K</td>
</tr>
<tr>
<td>3</td>
<td>16</td>
<td>XRE07K</td>
</tr>
<tr>
<td>4</td>
<td>39</td>
<td>ZWR100</td>
</tr>
<tr>
<td>4</td>
<td>40</td>
<td>ZWR110</td>
</tr>
</tbody>
</table>

3. The component has failed and is critical to system failure.\(^{24}\)

If failure of a min cut is confirmed while checking the \(n\)'th component, then outcome 3 must be true for that component.

\(^{24}\) Recall from Sec. II that a component is critical to system failure if \(\psi(1, Y(t)) - \psi(0, Y(t)) = 1\).

When the \(n\)'th component is found to be failed, it is possible that more than one min cut sets containing this component are failed. At this point, we introduce the concept of a critical cut set that enables us to handle this situation. First, as an example consider the reliability network diagram given below,

![Reliability Network Diagram]

where nodes represent the components of some system. If the system fails in time according to the sequence of component failures 2, 3, 1, then component 1 is critical at the time of system failure and two min cut sets \((1, 2)\) and \((1, 3)\) are failed. The set of components \((1, 2, 3)\) is called a critical cut set for component 1, since each cut set in the set \((1, 2, 3)\) contains a min cut set that contains component 1. In general, for a set of

[Note: The diagram is a reliability network with nodes 1, 2, and 3 connected in series.]
events to be a critical cut set for event i, each min cut set in this set must contain event i.

We now set up an expression for the expected time required for system diagnosis to confirm that a critical cut set has occurred. First, we consider notation.

**IV.A. Notation**

We adopt the notation of Sec. III. In addition, let \( T_i \) denote the time required to check component \( i \); let \( q_i(t) = q_i; p_i = 1 - q_i \); let \( T_S \) = time to diagnose system failure; let \( (l^k, \sigma_{n-k}, Y^{n-n}) \) be the state vector of a system of \( N \) components from which \( n \) components have been checked \((n = N)\), \( k \) components have been found to be failed, and \( n - k \) components are not failed. Let \( C'(Y) \) denote the set of components that have not been checked and \( C'(Y) \) the set of components that have not been checked yet.

**IV.B. Expression to Minimize Checking Time**

An expression for the expected time to diagnose system failure, \( E(T_5) \), involves \( \sum_{i=1}^{N} 2^i \) terms, where \( i \) is the checking order. The first seven terms of \( E(T_5) \) according to checking order or stage are given by

\[
E(T_5) = T_1 + \left( T_2 p_1 g_2(0_1, 0) + T_3 p_1 q_2 g_2(0_1, 0_2, 0) + T_3 q_1 p_2 g_2(0_1, 1_2, 0) + T_3 q_1 q_2 g_2(0_1, 1_2, 1) \right),
\]

where the terms following the vertical braces are summed. Recall that

\[
\Delta g_i(\mathbf{q}) \overset{\text{def}}{=} g(1_i, \mathbf{q}) - g(0_i, \mathbf{q}) .
\]

Order refers to the number of components that have been checked.

The terms listed above have the following meaning: The first-order term reflects the fact that at least one component in the system must be checked to confirm that a critical cut set has occurred. Second-order terms appear when the occurrence of a critical cut set on the first step has not been observed. The product, \( p_1 \Delta g_2(0_1, \mathbf{q}) \), in the second-order term represents the probability that the first component to be checked was not failed and the second component to be checked is critical to system failure when the first component has not failed. The corresponding product in the other second-order term has a similar probabilistic meaning with regards to the criticality of the second component except that now the first component was observed to be failed. Third-order terms appear when occurrence of a critical cut set on the second step has not been observed. Once we observe the occurrence of a critical cut set, say on the \( n \)th step, higher-order terms drop out since the criticality expression, \( \Delta g_i(\mathbf{q}) \), in the \( (n + 1) \)th and higher terms is zero.

There are \( N! \) possible expressions for and values of \( E(T_5) \) according to the checking order of the various components. Note, however, that in the ordering given above, if we check component 2 first and component 1 second, the terms involving \( T_5, \ldots, T_n \) do not change. To determine which component to check first, we propose to minimize \( E(T_5) \) with respect to the first two terms and neglect the third- and higher-order terms.

\[
T_i + \left( T_i p_1 \Delta g_i(0_1, 0) + T_i q_1 \Delta g_i(1_1, 0_1, 0) \right) > T_i + \left( T_i p_1 \Delta g_i(0_1, 1) + T_i q_1 \Delta g_i(1_1, 1_1, 0) \right), \tag{2}
\]

for all \( i \neq \mathbf{0} \), then component \( i \) should be checked first. This reduces the number of expressions for \( E(T_5) \) from \( N! \) to \( N(N - 1)/2 \).

The argument can be extended each time we check a component in the system. In general, if we have checked \( n \) components in the system, the next component we should check again is determined by an expression similar to Eq. (2),

\[
T_i + \left( T_i p_1 \Delta g_i(0_1, 1^k, 0^{n-k}, Y^{n-n}) + T_i q_1 \Delta g_i(1_1, 1^k, 0^{n-k}, Y^{n-n}) \right) > T_i + \left( T_i p_1 \Delta g_i(0_1, 1^k, 0^{n-k}, Y^{n-n}) + T_i q_1 \Delta g_i(1_1, 1^k, 0^{n-k}, Y^{n-n}) \right), \tag{3}
\]

where \( i \) and \( j \in C'(Y) \). The optimization procedure implied by the expression, Eq. (3), is referred to in decision theory as a one-step-ahead optimization policy.\(^{25}\)

If, on the \( n \)th step, we observe that a critical cut set, say \( K_i \), has failed, we still must somehow determine if \( K_j \) is the only critical cut set failed. If it is not, then in the expression, Eq. (3), we set \( Y_i = 0 \) for all \( i \in K_j \) and continue to check for another failed critical cut set.

**IV.C. Series System**

We use the expression, Eq. (2), to determine which component should be checked first for a series system with \( N \) components. Such a system will fail if any component fails. In this case,

\[
g(\mathbf{q}) = 1 - \prod_{k=1}^{N} (1 - q_k) = 1 - \prod_{k=1}^{N} p_k ,
\]

\[
\Delta g_i(\mathbf{q}) = \prod_{k=1}^{N} p_k ,
\]

and Eq. (2) becomes

This implies that

\[ T_j + T_i \prod_{k \neq i} p_k > T_i + T_j q_i \prod_{k \neq i} p_k . \]

Since for reliable system \( \prod_{k} p_k \approx 1 \), we obtain

\[ T_j + T_i p_i > T_i + T_j p_i \]

or

\[ \frac{T_j}{q_i} > \frac{T_i}{q_i} \quad \text{for all } j \neq i . \]

The above inequality states that the component with the minimum value of \( T_i/q_i \) should be checked first, an intuitive result for a series system. Note that in the series case, the above result is globally optimal since third-order terms and higher are zero.

**IV.D. Parallel System**

For a parallel system of \( N \) components that will fail if all components fail,

\[ g(q) = \prod_{i=1}^{N} q_i, \]

\[ \Delta g_i(q) = \prod_{j \neq i} q_i, \]

and Eq. (2) becomes

\[ T_j + T_i q_i \prod_{i \neq j} q_i > T_i + T_j q_i \prod_{i \neq j} q_i, \]

or

\[ T_j \left(1 - \prod_{i \neq j} q_i \right) > T_i \left(1 - \prod_{i \neq j} q_i \right). \]

Again, for reliable system

\[ \prod_{i \neq j} q_i \approx 0, \prod_{i \neq j} q_i \approx 0 \]

which implies

\[ T_j > T_i \quad \text{for all } j \neq i . \]

Thus, according to this inequality for a parallel system, the component with the minimum checking time should be checked first, again an intuitive result.

The obvious disadvantage to the above scheme is that it minimizes \( E(T_s) \) with respect to the first two terms only. In a rigorous analysis, third- and higher order terms should also be considered in finding the true optimal checking order. However, it is conjectured that it is extremely difficult and time consuming to set up a generalized expression that minimizes \( E(T_s) \). On the contrary, the expression, Eq. (2), is easy to compute and gives intuitively verifiable results for the series and parallel cases.

If the fault tree on which the checking scheme is based includes all system failure modes, then all the failed min cuts will be found using the above scheme. However, some components may still be failed after finding all the failed min cut sets. Hence, the above scheme is suited for diagnosis under emergency conditions, where the immediate goal is to return the system to the unfailed state in the shortest time possible.

**V. DECISIONS REGARDING SYSTEM OPERATION WHEN SYSTEM FAULT CONDITIONS OCCUR**

Following identification by the operator of the basic events such as hardware failures and maintenance faults that have occurred, the increased risk of operating the system can be determined by quantitatively reevaluating the fault tree for the entire system. On the basis of such factors as the length of time it may take to repair components or to rectify human errors, or the hazard associated with loss of subsystems or components, decisions may be made regarding the continuing operation of the system. Basically, four choices are available to an operator when system fault conditions occur:

1. shut the system down and repair failed components
2. change the mode of system operation while repairing the components
3. operate the system and simultaneously repair it
4. operate the system without immediate repair.

For example, all four choices are, in principle, available to the operator of a nuclear power plant if an engineered safeguard system is found inoperable. Choices 1 and 4 are available to a pilot who finds a hydraulic system inoperable in flight; i.e., he may land his aircraft at the nearest airport or continue his flight to his final destination. The choice the operator makes is influenced basically by the costs and hazards associated with system shutdown and restart versus the costs and hazards of continuing to operate at a higher risk level. For example, hazards and economic penalties associated with shutdown in the chemical industry are numerous: When hydrocarbon plants are shut down, the equipment must be purged to remove all oxygen before restart to avoid the
potential for fire. When shut down, stirred chemical reactors may develop hot spots if not thoroughly stirred, and explosions may result.

When nuclear power plants are shut down, thermal transients are induced, causing increased stresses on piping and equipment. Shutdown and restart also involve system transients that might increase the potential for accidents. Shutdown of these complex systems also results in lost production time and lost revenue. From this brief discussion, it is obvious that a trade-off analysis should be conducted when considering shutdown of highly complex and potentially hazardous systems.

V.A. Shutdown Decision at a Nuclear Power Plant

As an example of a shutdown decision made on a risk-assessment basis, consider a failure of low-pressure injection pump A in the example from Sec. III. Assume that this failure was revealed during a routine monthly test. The operator would like to know if this failure warrants plant shutdown. Technical specifications by the U.S. Nuclear Regulatory Commission (NRC) require the plant to be shut down to a hot standby condition if repair time, \( T \), is longer than 24 h, and to a cold standby condition if \( T \) > 48 h. The effect of the failure of pump A is that leg A is incapacitated until pump A can be repaired, and results in loss of LPIS redundancy. If a primary coolant system pipe rupture were to occur and if the leg-B pump failed to start, the potential for a large radiological release is increased. However, there is also a finite risk associated with plant shutdown.\(^{26}\) In Sec. V.A.1, we use the quantitative information presented in the Reactor Safety Study and in Table II of this paper to compute the risk of shutting down the plant versus the risk of plant operation with one LPIS pump out of service. The effect of thermal transients induced by shutdown and startup is included in this analysis. We then determine the repair time interval, \( T \), for which the risk associated with continuing plant operation becomes comparable to the risk of shutting down the plant. This analysis is necessarily of a simplified nature and is included here for illustration purposes only.

V.A.1. Establishing Maximum Allowable Repair Time, \( T \)

From Table II we see that with one LPIS pump out of service, the probability that the entire LPIS fails on demand is \( 7.9 \times 10^{-3} \), while the probability of both LPIS legs failing is \( 5.0 \times 10^{-4} \)

\(^{26}\)Here, risk is defined for simplicity only in terms of a core meltdown.

per demand. In case of a large pipe break, failure of the LPIS leads to core meltdown. For small pipe breaks, the high-pressure injection system can provide sufficient emergency cooling. The probability of a large pipe break is obtained from the Reactor Safety Study\(^{7}\) as \( 10^{-3}/\text{yr} \). Thus, the incremental hourly risk associated with plant operation with one LPIS pump out of service is

\[
\text{Prob (core meltdown/h | one LPIS pump failure)} = \text{Prob (large pipe break/h | LPIS failure)}
\]

\[
= \frac{10^{-4}/\text{yr} \times (1 \text{ yr}/8760 \text{ h})}{(7.9 \times 10^{-3} - 5.0 \times 10^{-4})} = 8.4 \times 10^{-11}/\text{h}
\]

(4)

The Reactor Safety Study considered PWR accident chains with loss of off-site power as the initiating event and concluded that these accident sequences contributed significantly to the overall risk of plant operation. If both the main feedwater and auxiliary feedwater systems fail to operate following this transient, the heat sink for decay heat removal is lost. The steam generators would be emptied in \( \frac{1}{2} \) h, causing the reactor coolant in the primary loop to heat up. The reactor coolant would be discharged through the pressurizer relief valves, causing the reactor core to be uncovered. Within \( \frac{1}{3} \) h after the transient, core melting would start. Various accident sequences were hypothesized that would result in loss of the main feedwater and auxiliary feedwater and auxiliary feedwater systems with loss of off-site power as the initiating event. As shown in Table IV, these sequences make a significant contribution to the probabilities of release across the entire release spectrum. For our example, we consider an orderly shutdown of the plant during which an operator error is committed that causes a turbine trip, which in turn imposes a transient instability in the electrical grid resulting in loss of off-site power. We estimate that the probability of operator error during shutdown causing a turbine trip is \( 10^{-2} \). Based on Federal Power Commission data, the probability of an off-site power loss during a turbine trip is \( 10^{-3} \) (Ref. 7). Thus, the probability of a core meltdown caused by loss of off-site power due to an operator error during shutdown is obtained by multiplying \( P_{TE,R} \) in Table IV by the ratio

\[
\frac{10^{-2} \times 10^{-3}}{0.2} = 5 \times 10^{-5}
\]

This probability per shutdown is

\[
5 \times 10^{-5} \sum_{R=1}^{7} P_{TE,R} = 8.0 \times 10^{-12}
\]

(5)
The Relevance of Multiattribute Utility Theory from Operation with a Partly Failed Unimpaired but Different from the Value During Steady-State Plant Startup Begins. The After Repair or Replacement of the LPIS: Values of Maximum Permissible Repair Times Versus Alpha

<table>
<thead>
<tr>
<th>Release Category $R$</th>
<th>$P_{TE,R}(\text{yr}^{-1})$</th>
<th>$P_{TE,R}/R_{\text{total},R} \times 100%$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$3 \times 10^{-7}$</td>
<td>38</td>
</tr>
<tr>
<td>2</td>
<td>$3 \times 10^{-6}$</td>
<td>41</td>
</tr>
<tr>
<td>3</td>
<td>$4 \times 10^{-7}$</td>
<td>10</td>
</tr>
<tr>
<td>4</td>
<td>$7 \times 10^{-8}$</td>
<td>15</td>
</tr>
<tr>
<td>5</td>
<td>$2 \times 10^{-7}$</td>
<td>32</td>
</tr>
<tr>
<td>6</td>
<td>$2 \times 10^{-6}$</td>
<td>41</td>
</tr>
<tr>
<td>7</td>
<td>$1 \times 10^{-5}$</td>
<td>26</td>
</tr>
</tbody>
</table>

*From Ref. 7:

$P_{\text{total},R} =$ probability of release of a category $R$ per year from all causes

$P_{TE,R} =$ probability of release of category $R$ due to loss of off-site power as an initiating event.

The probability of the transient-event accident sequence, $P_{TE}$, took the general form

$$P_{TE} = P_1 \prod_{i=2}^{n} P_i,$$

where

$P_1 =$ probability of loss of off-site power during normal operation

$= 0.2 \text{ occurrence/yr}$

$P_i =$ probability of occurrence of the $i$'th event in the accident sequence.

During shutdown and startup, thermal transients may also increase the hourly probability of a LOCA. We assume that the probability of a LOCA is the same during shutdown and startup, but different from the value during steady-state generation. We define the probability ratio, $\alpha$, as

$$\alpha = \frac{\text{Prob (LOCA during shutdown or startup/h)}}{\text{Prob (LOCA at steady state/h)}}.$$

After repair or replacement of the LPIS pump, plant startup begins. The LPIS unavailability is given in Table II as $5.0 \times 10^{-4}$. We estimate that the time required for plant startup and shutdown is 12 h for each. Then, the increased risk due to shutdown and startup includes the contributions from operation with a partly failed LPIS and an unimpaired LPIS:

$$\alpha(10^{-4}/\text{yr})(1 \text{ yr/8760 h})(7.9 \times 10^{-3} + 5.0 \times 10^{-4})$$

$$= 12 \text{ h} - (10^{-4}/\text{yr})(1 \text{ yr/8760 h})(5.0 \times 10^{-4}) 24 \text{ h}$$

$$= 1.2 \times 10^{-9} \alpha - 1.4 \times 10^{-10}.$$

(6)

We have not considered possible failure of other ECCSs and small pipe breaks to simplify the analysis. We obtain the allowed repair time, $T$, as a function of $\alpha$ by equating the incremental risks associated with the two options from Eqs. (4), (5), and (6).

8.4 $\times 10^{-11}$

$$= 8.0 \times 10^{-10} + 1.2 \times 10^{-9} \alpha - 1.4 \times 10^{-10}$$

or

$$\alpha = 0.07 \ T - 0.55.$$ (7)

Some values of the maximum permissible repair time versus $\alpha$ from Eq. (7) are given in Table V.

We see that if $\alpha = 1$, $T$ is given as 22 h, a lower bound in this analysis. The actual value of $\alpha$ dictates the value for $T$. The above analysis is not intended to be rigorous. However, it does show how decisions regarding system operation based on a risk assessment can be made when system fault conditions occur.

Polk performed a similar analysis based on economic penalty. He considered the NRC regulation requiring PWR plant shutdown 2 h after the loss of one of the two redundant emergency de-power buses. He computed the cost of lost revenue while the plant is shut down and compared it to the increased expected cost of an accident involving radiological releases that might occur during continuing plant operation while repair or replacement is taking place. Thus, he has determined the optimal allowable repair time based on cost.

Expected cost or increased risk may not be the only consequences considered in a trade-off analysis. A multiplicity of values involving safety, ecology, cost, availability of electrical power, etc. is involved in the process of imposing regulations on nuclear power plants. This suggests the use of multi-attribute utility theory in the decision process. Such an attempt is, however, clearly beyond the scope of this paper.

### TABLE V

<table>
<thead>
<tr>
<th>Values of Maximum Permissible Repair Times Versus Alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
</tr>
<tr>
<td>$T$ (h)</td>
</tr>
</tbody>
</table>

27 The unavailability of the LPIS with one pump out of service should dominate all other engineered safeguard system unavailabilities, making the expression, Eq. (6), an accurate approximation.


VI. CONCLUSION

The checklist generation scheme presented here provides an efficient way of diagnosing system malfunctions for complex systems. If the times required to check components vary widely, we may consider implementing the one-step-ahead optimization procedure on a computer. This is particularly desirable for potentially complex and hazardous systems where time available for diagnostic checking might be limited. The advantage to checklists is that an on-line computer is not needed since those can be prepared in advance for possible system malfunctions. In either method, we use the information gained as the system is examined to determine what component should be checked next.

When shutdown of a complex and hazardous system is considered following occurrence of system fault conditions, we should perform a trade-off analysis based on cost, safety, and other factors to actually determine if shutdown is the optimal course of system operation.

ACKNOWLEDGMENTS

The material for this article was extracted from Ref. 13, the PhD thesis of the first author. The authors wish to thank Richard Barlow of the University of California, Berkeley, for his valuable guidance of the research.

This work was performed under the auspices of the U.S. Energy Research and Development Administration.