NUREG/CR-1860
SAND80-2624
AN

ON THE RELATION OF VARIOUS RELIABILITY MEASURES TO EACH OTHER AND TO GAME THEORETIC VALUES*

Richard Engelbrecht-Wiggans
University of Illinois
Department of Mechanical and Industrial Engineering

and

David R. Strip
Fuel Cycle Risk Analysis
Division 4413

Published: January 1981

Sandia National Laboratories
Albuquerque, New Mexico 87185
operated by
Sandia Corporation
for the
U.S. Department of Energy

Prepared for
Division of Safeguards, Fuel Cycle & Environmental Research
Office of Nuclear Regulatory Research
U.S. Nuclear Regulatory Commission
Washington, D.C. 20555
Under Memorandum of Understanding DOE 40-550-75
NRC FIN No. A-1060

*This work supported in part by the National Science Foundation under Contract SOC 78-25219.

3-4

8105120099

We wish to acknowledge helpful discussions with Robert J. Weber, who, among other things, suggested the numerical example used in this paper.

Abstract

A variety of measures have recently been proposed for measuring the relative importance of individual components in the overall reliability of a system. Several of these seemingly different measures are very closely related under the conditions typically assumed in the reliability literature. The measures are also closely related to the probabilistic values of game theory.

1. Introduction

In many systems, the effect on the system of the failure of a given component is dependent on which other components have previously failed. For example, in nuclear safety systems a high degree of redundancy is provided, and therefore very few components can cause system failure in the absence of other component failures. Because of the complex relationships introduced through this redundancy, it becomes difficult to assess the relative contribution, or relative importance, of each component to the overall reliability of the system. This importance, however, would be a useful quantity to know when designing, modifying, or protecting a system subject to failure.

A number of seemingly different measures for the importance of individual components have been proposed; these include those of Barlow and Proschan,[1] Birnbaum,[2] and Fussell.[3] Although these measures have, on occasion, been discussed together (e.g., Lambert[4]), their relationships have not been developed. We will define a broad class of measures which include those mentioned, and examine their properties under specified conditions.

The above reliability measures are closely related to the concept of probabilistic values in game theory; probabilistic values measure the relative contributions of the players to the outcome of the game. Although the work by Banzhaf[5] and Shapley[6] predates the corresponding reliability literature by up to two decades, the connection has apparently not been observed before now. We will use certain game theoretic results, translated into the context of reliability, to critically compare a variety of measures for the importance of individual components in the reliability of a system.

2. Model

Let $N = \{1, 2, ..., n\}$ denote the set of components comprising a system, while S_t denotes the subset of components not functioning at time t. S_t^c is the complement of S_t, and therefore the set of components functioning at time t. The system is defined by probability functions $P_t(S)$ and $Q(S)$. The probability $P_t(S)$ that $S_t = S$ can include the possibilities that failed components are eventually repaired, and that each component's failure rate depends on time and on the status of the other components; $Q(S)$ is the probability that the system is not functioning if $S_t = S$. The probability $H(t)$ that the system is not functioning at time t is thus given by the expression $\Sigma_S\ Q(S)\ P_t(S)$. Finally, define $\mathscr{P}_X(t) = \Sigma_{S:X\subseteq S}\ P_t(S)$; the (marginal) probability that each of the elements of X is not functioning at time t.

Systems without the possibility of repair are often described by the joint probability distribution of the failure times of the individual components. Although computing $P_t(S)$ from such distributions may be non-trivial, it is clear that, at least in theory, the $P_t(S)$ may be obtained from the joint distribution of failure times. Although such notation is not commonly used, describing the failure times via the functions $P_t(S)$ simplifies our exposition.

The function Q(S) typically takes only values of zero or one; the value one corresponds to a set "critical" for the functioning of the system. This zero-one case is the structure function of Birnbaum.[2] More general functions allow non-trivial (i.e. not either zero or one) probabilities that certain sets are critical. For example, there may be a subset of components in a nuclear power facility such that if all these components are non-functional the system could still function; however, there may be a non-trivial probability that the operator will order the facility shut down due to factors external to the model. The function Q(S) is assumed to be monotonic in the sense that if S is contained in S', then Q(S) is no greater than Q(S'). For the zero-one case, this monotonicity is called coherence by Barlow and Proschan.[1]

The following two assumptions on $P_t(S)$ are frequently made in the literature on reliability; we will not make these assumptions unless specifically stated.

Assumption 1 (Independence of Failures):

$$P_t(S) = \prod_{i \in S} \mathscr{P}_i(t) \prod_{i \notin S} (1-\mathscr{P}_i(t)),$$

where $\mathscr{P}_i(t)$ is shorthand for $\mathscr{P}_{\{i\}}(t)$.

Assumption 2 (Symmetry of Failures): $P_t(S)$ is a function solely of $|S|$ (i.e., the number of components in S); components are interchangeable for purposes of computing $P_t(S)$.

3. Importance Measures

The first class of measures considered are of the form "what is the probability that the system status changes when the characteristics of a particular subset of components are altered?" Although the measures are frequently applied to single components, the reliability of sets will become important in a later section when we discuss modules. We will express these measures in two ways: once directly in terms of the functions $P_t(\cdot)$ and $Q(\cdot)$, and once as probabilistic expressions. Although these measures are typically defined for individual components in systems with each Q(S) having values zero or one, the definitions are easily stated in a more general form. Thus, consider the following importance measures.

$$M_{1,t}(X) = \Sigma_S \, P_t(S) \, [Q(S)-Q(S\setminus X)]$$

$$M_{2,t}(X) = \Sigma_S \, P_t(S) \, [Q(S\cup X)-Q(S)]$$

$$M_{3,t}(X) = \Sigma_S \, P_t(S) \, [Q(S\cup X)-Q(S\setminus X)] = M_{1,t}(X) + M_{2,t}(X)$$

$M_{4,t}(X) = dH(t)/d\mathscr{P}_X(t)$. (This is defined only if H(t) is a function of $\mathscr{P}_X(t)$).

$$M'_{1,t}(X) = \Pr\{\text{System Failed}\} - \Pr\{\text{System Failed} \mid X \subseteq S_t^c\}$$

$$M'_{2,t}(X) = \Pr\{\text{System Failed} \mid X \subseteq S_t\} - \Pr\{\text{System Failed}\}$$

$$M'_{3,t}(X) = \Pr\{\text{System Failed} \mid X \subseteq S_t\} - \Pr\{\text{System Failed} \mid X \subseteq S_t^c\} = M'_{1,t} + M'_{2,t}$$

Measures $M_{1,t}$, $M'_{1,t}$ may be interpreted as the probability that the system status at time t changes if all the non-functioning components in X are repaired at time t. Alternatively, it is the change in the probability that the system functions at time t when the components in the set X are made fail-proof. This appears an appropriate measure for those defending a system against attack or trying to improve system reliability by selectively upgrading some components. Measures $M_{2,t}$ and $M'_{2,t}$ may be interpreted as the probability that the system status at time t changes if all the functioning components in X are broken at time t. Such a measure might be of interest to anyone planning to sabotage a system. Finally, measures $M_{3,t}$, $M'_{3,t}$, and $M_{4,t}$ may be interpreted as the sensitivity of the reliability of the system to whether or not all the components of the set X are functioning at time t.

Although the above measures have been stated as reliability measures, measure $M_{1,t}$ is closely related to the game theoretic "probabilistic values," v_i ($i=1, 2, \ldots, n$), defined by $v_i = \Sigma_S P(S) [Q(S \cup i)-Q(S)]$, where the P(S)'s may be any non-negative numbers summing to one. A game theoretic question analogous to that of reliability is: given the value of each coalition of players, what is the contribution of any one player to the overall value. For example, in a voting situation, the value of a coalition is zero or one; a coalition's value is one if and only if the members of the coalition, working together, can assure the passage of a bill. The importance of a particular individual depends on how many votes he has (or controls) and how often his votes can influence the outcome.

Shapley[6] first introduced the notion for the value of a game by providing a set of axioms and deriving a unique set of P(S)'s which produce a value satisfying these axioms. Under the assumptions of independence and symmetry of failures, the structural measure of Barlow and Proschan,[1] which is defined only for binary Q(•), has $P_t(S)$'s identical to those of Shapley. The measure introduced by Banzhaf[5] also has this general form, but uses a different set of $P_t(S)$.

In order to compare Fussell's measure, an additional definition is needed. As in other reliability work, Fussell assumes Q(•) binary. A set S is called a minimal cut set if Q(S) = 1, and Q(S') = 0 for all S' contained in S. Fussell's measure for a component i is then Pr $\{$at least one minimal cut set containing i is failed | system is failed$\}$. This measure is similar to $M_{1,t}$ and $M'_{1,t}$, and attempts to capture the idea that a component i is <u>contributing</u> to failure. The main difference between Fussell's measure and $M_{1,t}$ (other than scaling) is that $M_{1,t}$ not only looks for involvement in a failed cut set, but requires that repair of i causes restoration of system function. Thus, i must be in all failed cut sets to be considered critical under $M_{1,t}$. Fussell's measure was defined with safety systems in mind, and therefore, the P(S) are such that the probability of two minimal cut sets existing simultaneously is negligible, and therefore the measure is virtually identical to $M_{1,t}$.

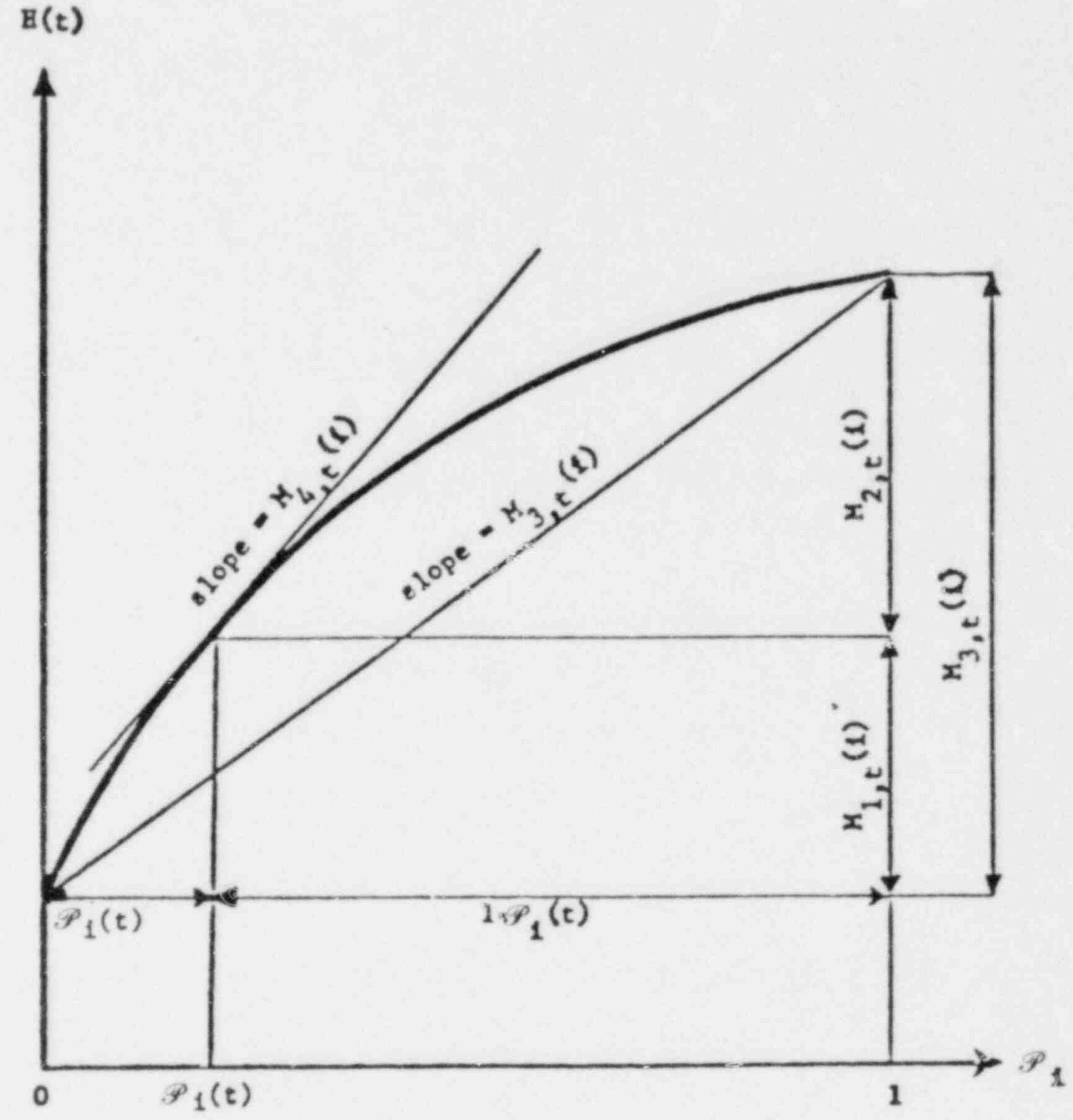

Figure 1

Relation of Measures for General System Reliability Function

Work in other areas fits within the general definition of the measures given here. Among these is Chow's[7] work on characterization of threshold functions. Dubey and Shapley[8] present a number of other applications of the type of measures discussed here.

4. Relationships Among Measures

In order to relate the various reliability measures, it is useful to assume that $H(t)$ is a function of $\mathscr{P}_i(t)$ for each i. Although this assumption rules out many forms of dependencies among the failure distributions of different components, the assumption is consistent with many reliability models.

Measures $M_{1,t}$, $M_{2,t}$, $M_{3,t}$ and $M_{4,t}$ are illustrated in Figure 1. The first three measures correspond to intervals, while the fourth (and also the third) correspond to slopes. This figure, together with some results of Owen[9] on multilinear extensions of games, suggests the following theorem.

Theorem 1: When $H(t)$ is a function of $\mathscr{P}_i(t)$ $\forall i$, then the following four conditions are equivalent.

1. $P_t(S) = \prod_{i \in S} \mathscr{P}_i(t) \prod_{i \notin S} (1-\mathscr{P}_i(t))$ $\forall t$.

2. For any function Q, the corresponding H(t) is given by

$$H(t) = \Sigma_S\, Q(S) \prod_{i \in S} \mathscr{P}_i(t) \prod_{i \notin S} (1-\mathscr{P}_i(t)) \;\forall t.$$

3. For any function Q, the corresponding H(t) is linear in each $\mathscr{P}_i(t)$, and $H(t) = Q(S)$ whenever $\mathscr{P}_i(t) = 0 \ \forall i \in S$ and $\mathscr{P}_i(t) = 1 \ \forall i \notin S$.

4. For any function Q, the corresponding measures satisfy

$$M_{1,t}(i)/\mathscr{P}_i(t) = M_{2,t}(i)/(1-\mathscr{P}_i(t)) = M_{3,t}(i)$$
$$= M_{4,t}(i) \ \forall i,t.$$

Proof: It is clear that 1 => 2 => 3. It follows from Figure 1 that 3 <=> 4. Owen[9] proves that 3 => 2. Finally, 2 => 1 follows from comparing the second condition to the definition $H(t) = \Sigma_S \, Q(S) \, P_t(S)$ for all Q.

Condition 1 is the Independence of Failures assumption described in Section 2. Although this assumption is common to most of the literature on reliability measures, the relationship it implies is apparently not mentioned in the literature. Because the numerators of the relationship in 4 are dependent on i, the vectors $\underline{M}_{j,t} = (M_{j,t}(1), M_{j,t}(2), \ldots, M_{j,t}(n))$, (j=1,2,3,4) are not proportional, and therefore the relationship is not obvious from numerical results.

The relationship between measures $M_{j,t}$ and $M'_{j,t}$ is examined in Theorem 2. In order to prove the theorem, however, we begin with two lemmata. (Note: We assume that $0 < P_i(t) < 1$. If $\mathscr{P}_i(t) = 0$ or 1, the component is irrelevant to the system in the sense that it cannot change state, and therefore cannot cause a change in system state.)

Lemma 1. If $\frac{P_t(S \cup i)}{\mathscr{P}_i(t)} - \frac{P_t(S)}{1-\mathscr{P}_i(t)} = 0 \quad \forall i, \forall S, i \notin S$

then $P_t(S) = \prod_{i \in S} \mathscr{P}_i(t) \prod_{i \notin S} (1-\mathscr{P}_i(t)) \; \forall S$

Proof: Consider S, $i \in S$. By hypothesis

$$P_t(S) = \frac{P_t(S \setminus i)\mathscr{P}_i(t)}{1-\mathscr{P}_i(t)} = P_t(\emptyset) \prod_{j \in S} \frac{\mathscr{P}_j(t)}{1-\mathscr{P}_j(t)}$$

$$= \frac{P_t(\emptyset)}{\prod_{j \in N} (1-\mathscr{P}_j(t))} \left[\prod_{j \in S} \mathscr{P}_j(t) \times \prod_{j \notin S} (1-\mathscr{P}_j(t))\right]$$

Since $\sum_S P_t(S) = 1$ and $\sum_S [\prod_{i \in S} \mathscr{P}_i(t) \prod_{i \notin S} (1-\mathscr{P}_i(t))] = \prod_{i \in N} [\mathscr{P}_i(t)+(1-\mathscr{P}_i(t))] = 1$

Then $\frac{P_t(\emptyset)}{\prod_{i \in N} (1-\mathscr{P}_i(t))} = 1$

Therefore

$$P_t(S) = \prod_{i \in S} \mathscr{P}_i(t) \prod_{i \notin S} (1-\mathscr{P}_i(t))$$

Lemma 1′: $\frac{P_t(S \cup i)}{\mathscr{P}_i(t)} - \frac{P_t(S)}{1-\mathscr{P}_i(t)} = 0 \quad \forall i, \forall S, i \notin S \iff$ Assumption 1 holds (independence of failure)

Proof: => by Lemma 1; <= by laws of probability.

Theorem 2. $M_{3,t}(i) = M'_{3,t}(i)\ \forall i,\ \forall Q \Leftrightarrow$ Assumption 1 holds.

Proof: $$M'_{3,t}(i) = \sum_{\substack{S \\ i \in S}} \frac{P_t(S)}{\mathscr{P}_i(t)} Q(S) - \sum_{\substack{S \\ i \notin S}} \frac{P_t(S)}{1-\mathscr{P}_i(t)} Q(S)$$

$$= [(1-\mathscr{P}_i(t))+\mathscr{P}_i(t)] \sum_{\substack{S \\ i \in S}} \frac{P_t(S)}{\mathscr{P}_i(t)} Q(S) - [(1-\mathscr{P}_i(t)+\mathscr{P}_i(t)] \sum_{\substack{S \\ i \notin S}} \frac{P_t(S)}{1-\mathscr{P}_i(t)} Q(S)$$

$$= \sum_{\substack{S \\ i \in S}} P_t(S)Q(S) + \frac{1-\mathscr{P}_i(t)}{\mathscr{P}_i(t)} \sum_{\substack{S \\ i \in S}} P_t(S)Q(S) - \sum_{\substack{S \\ i \notin S}} P_t(S)Q(S) - \frac{\mathscr{P}_i(t)}{1-\mathscr{P}_i(t)} \sum_{\substack{S \\ i \notin S}} P_t(S)Q(S)$$

interchange and rewrite the 2nd and 4th terms

$$= \sum_{\substack{S \\ i \in S}} P_t(S)Q(S) - \sum_{\substack{S \\ i \in S}} \frac{\mathscr{P}_i(t)}{1-\mathscr{P}_i(t)} P_t(S\backslash i)Q(S\backslash i) - \sum_{\substack{S \\ i \notin S}} P_t(S)Q(S) + \sum_{\substack{S \\ i \notin S}} \frac{1-\mathscr{P}_i(t)}{\mathscr{P}_i(t)} P_t(S \cup i)Q(S \cup i)$$

$$= \sum_{\substack{S \\ i \in S}} [P_t(S)Q(S) - P_t(S)Q(S\backslash i) + P_t(S)Q(S\backslash i) - \frac{\mathscr{P}_i(t)}{1-\mathscr{P}_i(t)} P_t(S\backslash i)Q(S\backslash i)]$$

$$- \sum_{\substack{S \\ i \notin S}} [P_t(S)Q(S) - P_t(S)Q(S \cup i) + P_t(S)Q(S \cup i) - \frac{1-\mathscr{P}_i(t)}{\mathscr{P}_i(t)} P_t(S \cup i)Q(S \cup i)]$$

$$= \sum_{\substack{S \\ i \in S}} P_t(S)[Q(S)-Q(S\backslash i)] + \mathscr{P}_i(t) \sum_{\substack{S \\ i \in S}} Q(S\backslash i) \left[\frac{P_t(S)}{\mathscr{P}_i(t)} - \frac{P_t(S\backslash i)}{1-\mathscr{P}_i(t)}\right]$$

$$+ \sum_{\substack{S \\ i \notin S}} P_t(S)[Q(S \cup i)-Q(S)] - (1-\mathscr{P}_i(t)) \sum_{\substack{S \\ i \notin S}} Q(S \cup i) \left[\frac{P_t(S)}{1-\mathscr{P}_i(t)} - \frac{P_t(S \cup i)}{\mathscr{P}_i(t)}\right]$$

Combining 1st and 3rd terms

$$= \sum_{S} P_t(S)[Q(S \cup i) - Q(S \setminus i))] + \mathscr{P}_i(t) \sum_{\substack{S \\ i \notin S}} Q(S) \left[\frac{P_t(S \cup i)}{\mathscr{P}_i(t)} - \frac{P_t(S)}{1-\mathscr{P}_i(t)} \right] +$$

$$(1-\mathscr{P}_i(t)) \sum_{\substack{S \\ i \notin S}} Q(S \cup i) \left[\frac{P_t(S \cup i)}{\mathscr{P}_i(t)} - \frac{P_t(S)}{1-\mathscr{P}_i(t)} \right]$$

Terms 2 and 3 = 0 (for all non-trivial Q) <=> factors in [ ] are 0. (This can be seen by considering the family of Q's $\left\{ Q_T : Q_T(S) = 0,\ S \neq T,\ Q_T(S) = 1,\ S = T \right\}$. By lemma 1', this occurs <=> assumption 1 holds.

Therefore

$$M'_{3,t}(i) = \sum_{S} P_t(S)[Q(S \cup i) - Q(S \setminus i)] = M_{3,t}(i) \iff \text{assumption 1 holds.}$$

Corollary: for each j, j=1,2,3 $M_{j,t}(i) = M'_{j,t}(i)$ <=> assumption 1 holds.

Theorems 1 and 2 clearly show the strong structure induced on the measures by the independence of failures assumption. Since this assumption is the basis for much of the existing literature, these Theorems demonstrate the equivalence in computing the various measures. Continuing to use the independence assumption may severely restrict the ability to accurately model many situations, since failures in complex systems, such as nuclear power plants, often have a sequential or common cause basis which cannot be modeled under the independence assumption. One approach to modeling common cause failures is to "factor out" the common cause as a separate "component", and reduce the failure probabilities of the involved components by an appropriate amount. This results, however, in two mutually exclusive events, which are obviously not independent.

5. Causal Measures

An alternative measure for the importance of a particular component is the probability that the system fails due to the failing of this particular component. In such a measure, a component contributes to the fallibility of the system only when it is the proverbial "straw which breaks the camel's back." A component which always fails before the system fails (but is never the last component to fail before the system fails) contributes nothing to system failures and is assigned a zero weight in such causal measures.

Consider the following two causal measures.

$$M_{5,t}(i) = \sum_S \Pr\{S_t=S \mid i \text{ fails at time } t^+\} \; [Q(S\cup i)-Q(S)]$$

$$M_{6,t}(i) = \sum_S \Pr\{S_t=S \mid i \text{ repaired at time } t^+\} \; [Q(S)-Q(S\setminus i)],$$

where "i fails at time t^+" is interpreted as "$i \notin S_t$ but $S_{t+e} = S_t \cup i$ for all sufficiently small positive e," and "i repaired at time t^+" is to be interpreted as "$i \in S_t$, but $S_{t+e} = S_t \setminus i$ for all sufficiently small positive e." Note that this implicitly assumes zero probability of more than one failure and/or repair occurring simultaneously.

Under independence of failures, the probabilities in $M_{5,t}(i)$ simplify to $P_t(S)/(1-\mathscr{P}_i(t))$, while the probabilities in $M_{6,t}(i)$ simplify to $P_t(S)/\mathscr{P}_i(t)$. This observation, together with Theorem 1, yields the following result.

Theorem 3: If assumption 1 holds, then $M_{6,t}(i) = M_{5,t}(i) = M_{4,t}(i) = M_{3,t}(i) = M_{2,t}(i)/(1-\mathscr{P}_i(t)) = M_{1,t}(i)/\mathscr{P}_i(t)$.

Barlow and Proschan[1] consider systems without repair and define the importance M^* of component i as the probability that the failing of i causes the system to fail. Under independence of failures, the expected number of system failures caused by component i during the time interval T is $\int_{t\in T} M_{5,t}(i)\, d\mathscr{P}_i(t)^+$, where $d\mathscr{P}_i(t)^+$ denotes the positive part of $d\mathscr{P}_i(t)$. If there are no repairs, then this expectation is equal to M^*. Under independence of failures, Theorem 3 may be used to relate M^* to the sensitivity measures.

In models with no repairs and symmetry of failures, it is easy to verify that each of the n! possible orders in which the n components may fail are equally likely to occur. Thus, the probability that the failure of component i is preceeded by the failures of all the elements of S is $s!(n-s-1)!/n!$, where s denotes the number of elements in S.

Theorem 4: If there is symmetry of failures and there are no repairs, then

$$M^* = \sum_S [s!(n-s-1)!/n!][Q(S\cup i)-Q(S)]$$

The above value of M^* is also the game theoretic Shapley[6] value; the unique game theoretic value satisfying certain axioms, including one of symmetry.

If $M_{3,t}(i)$ or $M_{4,t}(i)$ is independent of t in a model without repairs and with symmetric and independent failures, then these measures are equal to $\sum_S [s!(n-s-1)!/n!][Q(S \cup i)-Q(S)]$. In particular, as Weber[10] observes for game theoretic probabilistic values, $M_{3,t}(i)$ and $M_{4,t}(i)$ cannot be equal to the Banzhaf[5] value (or, equivalently, the Birnbaum[2] measure evaluated at $F_i(t) = 1/2\ \forall i$) for all times t in models with symmetric, independent failures and no repairs.

Finally, an example shows that Theorem 4 need not hold if there are repairs. In particular, consider a three component system with $Q(\{1,2,3\}) = Q(\{1,3\}) = Q(\{2,3\}) = 1$, and $Q(S) = 0$ for all other S. Assume that the "uptimes" (time from completion of repair until the next failure) of the components are identical independent exponential random variables. Likewise, the "downtimes" (time from failure until completion of repair) are identical independent exponential random variables; the uptimes and downtimes are also assumed to be independent of each other. Thus, the failures and repairs are symmetric and independent.

Characterize the state of the system by the subset of failed components; there are eight possible states. The memoryless quality of exponential distributions results in a Markov process for the transitions from one state to another. This process is depicted graphically in Figure 2. The probability that the system goes from state i to state j conditional on the system being in state i are indicated in the figure as a function of p, where p is the fraction of time a component is failed, or more precisely, the mean downtime divided by the sum of the mean uptime and the mean downtime.

The above transition probabilities imply the following state probabilities: $P_t(S) = P_{|S|} \forall t$, where $P_0 = (1-p)^2/2$, $P_1 = (1+p)(1-p)/6$, $P_2 = p(2-p)/6$, and $P_3 = p^2/2$. Thus, the relative probabilities that the failing of component i causes the system to fail are the probabilities that the system goes from state {3} to state {1,3}, the probability that the system goes from state {3} to state {2,3}, and the probability that the system goes from state {1,2} to state {1,2,3}, from state {2} to state {2,3} or from state {1} to state {1,3}. Thus, it follows that the probability component i failing causes the system to fail, conditional on the system failing, is $(1-p)/(4-3p)$, $(1-p)/(4-3p)$, and $(2-p)/(4-3p)$, respectively, for i = 1, 2, and 3.

If p is close to one, then most of the system components are likely to be non-functional; under such cases, a system failure is most likely to have been caused by a transition from state $\{1,2\}$ to state $\{1,2,3\}$. Indeed, as p tends to one, the conditional probabilities that component i causes system failure tends to 0, 0, and 1 respectively. Alternatively, if p is very small, then the system is likely to be in a state with few failed components and system failures are likely to have been caused by any one of the three components failing. In particular, as p tends towards zero, the conditional probability that component i causes the system to fail tends to 1/4, 1/4, and 1/2 respectively. Note that in the above example with symmetric and independent failures and repairs, the relative importance of the three components (measured in terms of which component's failing causes the system to fail) depends on the parameter p and cannot satisfy the conclusion of Theorem 4 for arbitrary p. Thus, the assumption of no repairs is necessary for Theorem 4.

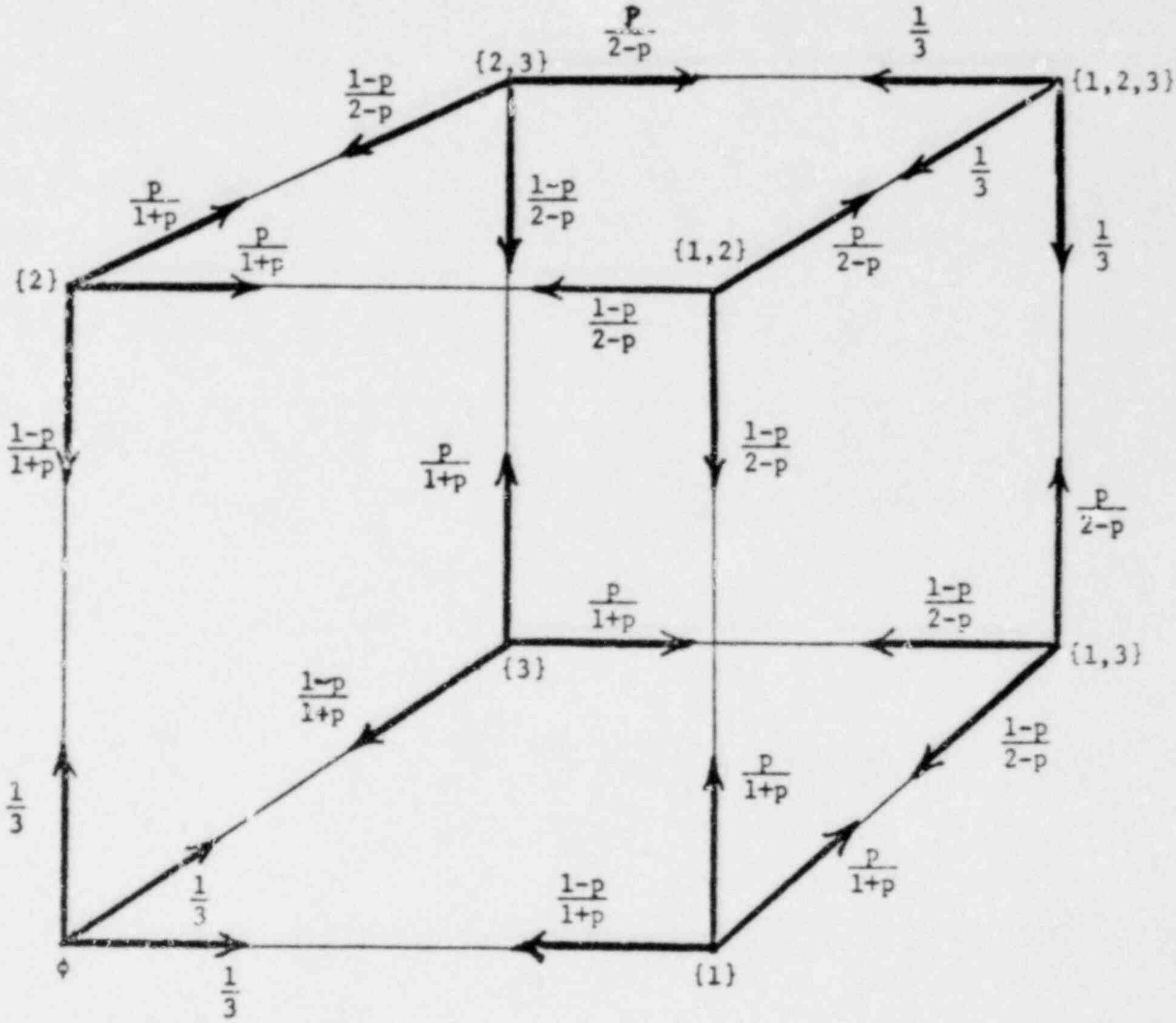

Figure 2

Markov Process Corresponding to Example

[illegible]. Modules

In many systems, there exist identifiable subsystems which can be treated as "super-components" in the sense that once one knows the state of the subsystem, knowing the state of the individual components within the subsystem adds no information in determining the state of the overall system. Birnbaum[2] develops this concept for binary Q functions, called structure functions. These "super-components" are called modules. For a system with component set N, and structure function Q (i.e., Q = 0 or 1), $\mathcal{M} \subseteq N$ is a module if there exist structure functions Q' and $Q_{\mathcal{M}}$ $Q': \{0,1\} \times N - \mathcal{M} \longrightarrow \{0,1\}$, $Q_{\mathcal{M}}: \mathcal{M} \cdot \{0,1\}$ and $Q(S) = Q'(Q_{\mathcal{M}}(S \cap \mathcal{M}), S \setminus \mathcal{M})$ $\forall S \subseteq N$. The notion of a module can be useful in calculating the various importance measures under certain conditions.

If $\mathcal{M}$ is a module of N, call $\mathcal{M}$ independent of $N-\mathcal{M}$ if $P(S) = P(S \cap \mathcal{M}) \cdot P(S \setminus \mathcal{M})$ $\forall S \subseteq N$. When modules can be found satisfying this condition, a form of 'chain rule' property exists which can simplify calculation, and expose more of the structure of the measures. In order to prove this property, we first start with a lemma.

Lemma: For a module $\mathcal{M}$, Q, Q', $Q_{\mathcal{M}}$ as above, $\forall i \in \mathcal{M}$, $\forall S \subseteq N$

a) $Q(S) - Q(S \setminus i) = [Q_{\mathcal{M}}(S \cap \mathcal{M}) - Q_{\mathcal{M}}(S \cap \mathcal{M} \setminus i)][Q'(1, S \setminus \mathcal{M}) - Q'(0, S \setminus \mathcal{M})]$

b) $Q(S \cup i) - Q(S) = [Q_{\mathcal{M}}(S \cap \mathcal{M} \cup i) - Q_{\mathcal{M}}(S \cap \mathcal{M})][Q'(1, S \setminus \mathcal{M}) - Q'(0, S \setminus \mathcal{M})]$

c) $Q(S \cup i) - Q(S \setminus i) = [Q_{\mathcal{M}}(S \cap \mathcal{M} \cup i) - Q_{\mathcal{M}}(S \cap \mathcal{M} \setminus i)][Q'(1, S \setminus \mathcal{M}) - Q'(0, S \setminus \mathcal{M})]$

Proof: (for form a))

Case I. $Q(S) - Q(S\backslash i) = 0$

$$Q_{\mathcal{M}}(S\cap\mathcal{M}) - Q_{\mathcal{M}}(S\cap\mathcal{M}\backslash i) = 0 \quad \text{(i.e., i has no effect on } \mathcal{M}\text{)}.$$

or

$$Q'(1,S\backslash\mathcal{M}) - Q'(0,S\backslash\mathcal{M}) = 0 \quad \text{(i.e., } \mathcal{M} \text{ has no effect on system)}.$$

Case II. $Q(S) - Q(S\backslash i) = 1$

$$Q_{\mathcal{M}}(S\cap M) - Q_{\mathcal{M}}(S\cap\mathcal{M}\backslash i) = 1 \quad \text{(i.e., i swings } \mathcal{M}\text{)}$$

and

$$Q'(1,S\backslash\mathcal{M}) - Q'(0,S\backslash\mathcal{M}) = 1 \quad \text{(i.e., } \mathcal{M} \text{ swings system)}.$$

b is a restatement of a with S replaced by $S\cup i$.

c is the sum of a and b.

<u>Theorem 5.</u> If $\mathcal{M}$, a module, is independent of $N\div\mathcal{M}$, then $M_{j,t}(i) = M_{3,t}(\mathcal{M})\ M_{j,t}(i)$ $i\in\mathcal{M}$, $j = 1,2,3$ (where $M_{3,t}(\mathcal{M})$ is the importance of $\mathcal{M}$ w.r.t. Q', and $M_{j,t}(i)$ is w.r.t. $Q_{\mathcal{M}}$).

Proof: For case j = 1

$$M_{1,t}(i) = \sum_{S\subseteq N} P(S)\ [Q(S) - Q(S\backslash i)]$$

$$= \sum_{S\subseteq N} P(S)\ [(Q_{\mathcal{M}}(S\cap\mathcal{M}) - Q_{\mathcal{M}}(S\cap\mathcal{M}\backslash i))(Q'(1,S\backslash\mathcal{M})-Q'(0,S\backslash\mathcal{M}))]$$

$$= \sum_{S\subseteq N} P(S\cap\mathcal{M})\ P(S\div\mathcal{M})[(Q_{\mathcal{M}}(S\cap\mathcal{M})-Q_{\mathcal{M}}(S\cap\mathcal{M}\backslash i))(Q'(1,S\backslash\mathcal{M}) - Q'(0,S\backslash\mathcal{M}))]$$

$$= \sum_{S\subseteq N\div\mathcal{M}} P(S)[Q'(1,S)-Q'(0,S)] \sum_{S\subseteq\mathcal{M}} P(S)[Q_{\mathcal{M}}(S)-Q_{\mathcal{M}}(S\backslash i)]$$

$$= \underset{\text{w.r.t.}Q'}{M_{3,t}(\mathcal{M})} \cdot \underset{\text{w.r.t.}Q_{\mathcal{M}}}{M_{1,t}(i)}$$

(The proofs for j = 2,3 are analogous.)

A similar result applies for measures $M'_{j,t}$, $j = 1,2,3$. However, since no structure function is defined for the models yielding measures $M'_{j,t}$, an alternative definition must be found. For these measures, $\mathcal{M}$ is a <u>probabilistic module</u> if, for all $i \in \mathcal{M}$

$$\Pr\{\text{System Failed}|\text{State of } \mathcal{M} \text{ and State of } i \in \mathcal{M}\} = \Pr\{\text{System Failed}|\text{State of } \mathcal{M}\}.$$

The chain rules for measures $M'_{j,t}$ are given in the next Theorem.

<u>Theorem 6</u>: If $\mathcal{M}$ is a probabilistic module, then for all $i \in \mathcal{M}$

$$\underset{}{M'_{j,t}(i)} = \underset{\text{w.r.t. system}}{M'_{3,t}(\mathcal{M})} \cdot \underset{\text{w.r.t. } \mathcal{M}}{M'_{j,t}(i)}.$$

Proof: Let X denote system failed,

$\overline{Y}$ denote module $\mathcal{M}$ failed,

Y denote module $\mathcal{M}$ functioning,

Z denote component i functioning.

$$\begin{aligned}M'_{1,t}(i) &= \Pr\{X\} - \Pr\{X|Z\} \\ &= \Pr\{X|Y\}\Pr\{Y\} + \Pr\{X|\overline{Y}\}\Pr\{\overline{Y}\} - \Pr\{X|Z \text{ and } Y\}\Pr\{Y|Z\} \\ &\quad - \Pr\{X|Z \text{ and } \overline{Y}\}\Pr\{\overline{Y}|Z\}.\end{aligned}$$

By modularity of $\mathcal{M}$

$$\begin{aligned}&= \Pr\{X|Y\}\Pr\{Y\} + \Pr\{X|\overline{Y}\}P\{\overline{Y}\} - \Pr\{X|Y\}\Pr\{Y|Z\} - \Pr\{X|\overline{Y}\}\Pr\{\overline{Y}|Z\} \\ &= \Pr\{X|Y\}[\Pr\{Y\} - \Pr\{Y|Z\}] + \Pr\{X|\overline{Y}\}[\Pr\{\overline{Y}\} - \Pr\{\overline{Y}|Z\}] \\ &= \Pr\{X|Y\}[\Pr\{Y\} - \Pr\{Y|Z\}] - \Pr\{X|\overline{Y}\}[\Pr\{Y\} - \Pr\{Y|Z\}] \\ &= [\Pr\{X|Y\} - \Pr\{X|\overline{Y}\}][\Pr\{Y\} - \Pr\{Y|Z\}] \\ &= M'_{3,t}(\mathcal{M}) \cdot M'_{1,t}(i)\end{aligned}$$

(Proofs for $j = 2,3$ are analogous).

Since the calculation of importance measures is, in general, exponential in difficulty, modules and chain rules can simplify computation significantly.

Conclusion

A number of apparently different measures for the importance of an individual component to the reliability of a system are examined in this paper. By defining all the measures within the same, sufficiently general, model, some insight is gained into the different probability questions corresponding to the different measures. It is, however, shown that under the (common) assumption of independence of failures, the importance measures are very closely related to each other. It is also shown that under the (common) assumption of symmetric failure rate distributions, the "causal" measures must take a particular form, and that they are also, over appropriate time intervals, equal to one of the previously mentioned sensitivity measures. Finally, the concept of modules is developed for the different types of measures, and a chain rule form is proven, which may be useful in computing the various measures.

REFERENCES

1. R. E. Barlow and F. Proschan, "Importance of System Components and Fault Tree Events," Stochastic Processes and Their Applications, Vol. 3, pp. 153-173 (1975).

2. Z. W. Birnbaum, "On the Importance of Components in a Multicomponent System" in P. R. Krishaiad, ed., Multivariate Analysis-II, Academic Press, New York (1969).

3. J. B. Fussell, "How to Hand Calculate System Reliability and Safety Characteristics," IEEE Transactions on Reliability, Vol. R24, #3 (August 1975).

4. H. E. Lambert, "Measures of Importance of Events and Cut Sets in Fault Trees," Reliability and Fault Tree Analysis, SIAM, pp. 77-100 (1975).

5. J. F. Banzhaf III, "Weighted Voting Doesn't Work: A Mathematical Analysis," Rutgers Law Review, Vol. 19, pp. 317-343 (1965).

6. L. Shapley, "A Value for n-Person Games," Annals of Mathematics Study, Vol. 28, pp. 307-318 (1953).

7. C. K. Chow, "On the Characterization of Threshold Functions," in Switching Theory and Logical Design, IEEE Special Publications S-134, pp. 34-38 (September 1961).

8. P. Dubey and L. S. Shapley, "Some Properties of the Banzhaf Power Index," Mathematics of Operations Research, Vol. 4, No. 2, (May 1979).

9. G. Owen, "Multilinear Extensions of Games," Management Science, Vol. 18, #5, Part II, pp. P64-P79 (1972).

10. R. J. Weber, "Subjectivity in the Valuation of Games," Cowles Foundation Discussion Paper #515, Yale University, New Haven, CT (January 1979).

DISTRIBUTION:

US NRC Distribution Contractor (CDSI) (130 copies for AN)
7300 Pearl Street
Bethesda, MD 20014

Richard Engelbrecht-Wiggans (10)
Department of Mechanical and
Industrial Engineering
226 MEB
University of Illinois
Urbana, IL 61801

David C. Heath
S.O.R.I.E.
Upson Hall
Cornell University
Ithaca, NY 14853

Major Peter Knepell
D5MS
United States Air Force Academy
Colorado

Howard Lambert
TERA Corporation
2150 Shattuck Avenue
Berkeley, CA 94704

400	C. Winter
1222	C. F. Diegert
1223	K. V. Diegert
1417	F. W. Spencer
4400	A. W. Snyder
4410	D. J. McCloskey
4413	N. R. Ortiz
4413	R. M. Cranwell
4413	D. R. Strip (10)
4414	G. B. Varnado
4414	R. B. Worrell
4416	L. D. Chapman
5641	P. J. Slater
5642	B. L. Hulme
5642	K. I. Hiebert
5642	R. J. Hanson
3141	T. L. Werner (5)
3151	W. L. Garner (3)
3154-3	J. Hernandez (25) for NRC distribution to NTIS
8214	M. A. Pound