KfK 3824
Dezember 1984

Methods of Fault Tree Analysis and Their Limits

G. G. Weber
Institut für Datenverarbeitung in der Technik

Kernforschungszentrum Karlsruhe

KERNFORSCHUNGSZENTRUM KARLSRUHE

Institut für Datenverarbeitung in der Technik

KfK 3824

METHODS OF FAULT TREE ANALYSIS AND THEIR LIMITS

by

G.G. Weber

Presented at the Conference on "Reliability Aspects: From Devices to Systems"
(Organized by the Materials and Testing Group of the Institute of Physics, UK,
11-12 September 1984, Brunel University, UK)

Kernforschungszentrum Karlsruhe GmbH, Karlsruhe

Kernforschungszentrum Karlsruhe GmbH
ISSN 0303-4003

Abstract

Fault tree analysis is a well known technique used for problems of system reliability. The subject of this paper is twofold:

- Some recent methodological developments of fault tree analysis will be discussed.
- Limits of fault tree analysis and a criterion for admissibility of structure functions will be given.

It will be shown that there are interesting relations to switching theory and to stochastic processes.

An introduction to some basic concepts and techniques of fault tree analysis will be given. We note that a fault tree can be defined as a directed graph. If we assume only two possible states per vertex, we obtain a Boolean function (structure function) which is equivalent to a combinational circuit. Such a system has the same configuration during its whole life. It is possible to evaluate unavailability and expected number of failures.

If we have, however, a system with a phased mission, its relevant configurations may change during consecutive periods (called phases). Systems which have to perform phased missions are, for instance, reactors with core cooling (during various phases of an accident) and fault tolerant aerospace computing systems (during various phases of a flight). Reliability and performance analysis requires the use of a (generalized) multistate structure function and the concept of association. It is possible to evaluate unavailability.

It is interesting to have here a criterion which can show the admissibility of phased structure functions for these systems. This is based on algebraic properties of functional dependence which again has strong relations to switching theory and to system analysis.

METHODEN DER FEHLERBAUMANALYSE UND IHRE GRENZEN

Zusammenfassung

Die Fehlerbaumanalyse ist eine bekannte Technik, die für Probleme der Systemzuverlässigkeit Verwendung findet. In der vorliegenden Arbeit werden zwei Themen behandelt:

- Einige neue methodische Entwicklungen der Fehlerbaumanalyse werden diskutiert.
- Grenzen der Fehlerbaumanalyse und ein Kriterium für die Zulässigkeit von Strukturfunktionen werden gezeigt.

Es stellt sich dabei heraus, daß interessante Beziehungen zur Schaltalgebra und zu stochastischen Prozessen bestehen.

Eine Einführung einiger grundlegender Begriffe und Techniken der Fehlerbaumanalyse wird gegeben. Wir stellen fest, daß ein Fehlerbaum als ein gerichteter Graph definiert werden kann. Nehmen wir an, daß jede Ecke des Graphen in nur zwei Zuständen sein kann, so erhalten wir eine Boole'sche Funktion (Strukturfunktion), die zu einem Schaltnetz äquivalent ist. Ein solches System hat dieselbe Konfiguration in seinem ganzen Leben. Es ist möglich, die Nichtverfügbarkeit und die erwartete Zahl der Ausfälle zu berechnen.

Haben wir jedoch ein System, das eine in Phasen aufgeteilte Mission ausführen soll, so können sich die Konfigurationen für aufeinanderfolgende Abschnitte der Mission ändern. Diese Abschnitte werden als Phasen bezeichnet. Systeme, die in Phasen aufgeteilte Missionen ausführen müssen, sind z.B. folgende:

Reaktoren mit Notkühlung (während verschiedenen Phasen eines Reaktorunfalls) sowie fehlertolerante Rechnersysteme für Flugzeuge (während verschiedenen Phasen eines Fluges). Die Analyse der Zuverlässigkeit und Leistungsfähigkeit erfordert die Verwendung einer (verallgemeinerten) Strukturfunktion mit mehrwertiger Logik sowie ein Verlassen des Bereichs der stochastischen Unabhängigkeit.

Es ist interessant, hier ein Kriterium zu haben, das die Zulässigkeit der Strukturfunktion für diese Systeme zeigen kann. Es basiert auf algebraischen Eigenschaften der "funktionalen Abhängigkeit", die wiederum stark mit Schaltalgebra und Systemanalyse verbunden sind.

Contents

1. Boolean Concepts of Fault Trees

A general introduction to fault tree analysis is given. Basic concepts of fault tree representation are introduced and relations to switching theory emphasized. The probabilistic evaluation of coherent systems is discussed. This is an application of alternating renewal processes. It is possible to use for evaluation minimal cuts, expansion, or modular decomposition. For decomposition, interesting relations to switching theory exist.

1.1. Definition and Representation of a Fault Tree

The subject of this chapter is a general introduction to fault tree analysis. The purpose of this analysis is twofold:

a) a systematic identification of all possible failure combinations which lead to a defined (undesired) event, i.e. system failure,

b) the evaluation of reliability and safety of a system (e.g. unavailability, unreliability, expected number of failures).

We shall not be concearned here with fault tree construction which is a very important step for modelling.

1.2 Definition of a Fault Tree

Although the term 'fault tree' is often used in a rather wide sense it seems preferable to us to concentrate on the following definition:

Definition

A fault tree is a finite directed graph without (directed) circuits. Each vertex may be in one of several states. For each vertex a function is given which specifies its state in terms of the states of its predecessors. The states of those vertices without predecessors are considered the independent variables of the fault tree /1/, /2/.

Some general properties of a fault tree:

1. The vertices without predecessors are the inputs to the fault tree, representing the components. We are interested in the state of every other vertex, but in particular with the state of one vertex without successors, an output vertex which we identify with the state of the system as a whole. The graphical term 'vertex' here is roughly synonymous with 'item' and generally denotes any level in the system, whether a component, sub-system or the whole system.

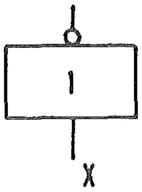
2. We mostly specialize to only two states per vertex. This makes all of the functions Boolean functions. We call one of the two states 'functioning', 'false' or 0, and the other 'failed', 'true' or 1.

OR, disjunction

$x_1 \vee x_2$ (1.2-1)

AND, conjunction

$x_1 \; x_2$ (1.2-2)



NOT, complement

$\bar{x}_1$ (1.2-3)

A few concepts related to Boolean functions

1. Coverage: A Boolean function $\Phi_1(\underline{x})$ is said to cover $\Phi_2(\underline{x})$, denoted $\Phi_1(\underline{x}) \supseteq \Phi_2(\underline{x})$, if Φ_1 assumes the value 1 whenever Φ_2 does.

2. Equivalence: If $\Phi_1 \supseteq \Phi_2$ and $\Phi_2 \supseteq \Phi_1$, Φ_1 and Φ_2 are equivalent.

3. Boolean monomial: A product term (monomial) is a conjunction $\prod_{i=1}^{n} x_i$ with no variable x_i complemented and uncomplemented.

4. Sum of products: A disjunction of Boolean monomials

$$\bigvee_{j=1}^{l} \prod_{i=1}^{n_j} x_{ij} \qquad (1.2\text{-}4)$$

is called a sum of products (sop) or polynomial.

5. Implicant: An implicant p_j of $\Phi(\underline{x})$ is a monomial which is covered by $\Phi(\underline{x})$.

6. Prime implicant: A prime implicant p_j of $\Phi(\underline{x})$ is an implicant which ceases to be an implicant if one variable is deleted from p_j.

Example: $p_j = \bar{x}_1 x_2$ is a prime implicant of the polynomial $\Phi(\underline{x}) = \bar{x}_1 x_2 + x_1 x_2 + x_2 \bar{x}_3$, but neither $\bar{x}_1$ nor x_2 alone implies $\Phi(\underline{x})$.

7. Base: A base of $\Phi(\underline{x})$ is a sop which is equivalent to $\Phi(\underline{x})$ where all monomials are prime implicants.

8. Irredundant base: A base which ceases to be a base if one prime implicant is deleted.

Remark: An irredundant base may be written

$$\Phi(\underline{x}) = \bigvee_{j=1}^{l} p_j \qquad (1.2\text{-}5)$$

where l is the number of prime implicants in the base, and p_j is the j^{th} prime implicant, given as

$$p_j = \prod_{i=1}^{n_j} x_{ij} \qquad (1.2-6)$$

9. Prime implicate: A dual set of concepts, leading to conjunctions of prime implicates. This will not be developed in detail but used if necessary /3/, /4/.

Note, that our definition of a two-state fault tree is equivalent to a combinational circuit with one output.

The no-circuit condition in the graph is equivalent to the condition that the current output of a switching circuit is entirely determined by current inputs, without memory of previous inputs or internal states.

Also the more general case of manyvalued logic and logic trees is included in this definition.

1.3 Boolean Approach

Structure function

We introduce the concept of structure function which is of central importance to fault tree analysis. It can be seen that it is closely related to the concept of switching function. We assume a system S, which has n components which can be in two states (functioning, failed). Also the system S can be in two states, either functioning or failed. The components are the vertices without predecessors of our fault tree definition. The function which specifies the state of a vertex in terms of its predecessors is a Boolean function (AND, OR, NOT). The states of the top vertex can be given by a structure function (see 1.2)/2/.

Definition of structure function

Let $x_1, x_2, \ldots, x_n$ be Boolean variables which can assume the values 0,1, where

$$x_i = \begin{cases} 0 \text{ if component } i \text{ is functioning} \\ 1 \text{ if component } i \text{ is failed.} \end{cases}$$

The assumption that 1 corresponds to failure is used throughout this paper and is useful for fault tree analysis. The Boolean variable x_i indicates the state of component i, whereas the state vector $\underline{x} = (x_1, x_2, \ldots, x_n)$ indicates the state of the system.

The Boolean function $\Phi(x_1, x_2, \ldots, x_n)$ is called structure function and determines completely the state of a system S in terms of the state-vectors:

$$\Phi(x_1, x_2, \ldots, x_n) = \begin{cases} 0 \text{ if system S is functioning} \\ 1 \text{ if system S is failed.} \end{cases}$$

Remark: The structure function is equivalent to a switching function representing a combinational circuit.

Combinational switching function

A combinational switching function is a mapping $f:B^n \rightarrow B$ where $B = \{0,1\}$ and B^n denotes the set of 2^n binary n-tuples. A switching function specifies for every input combination $(x_1, x_2, \ldots, x_n)$ an output value $y = f(x_1, x_2, \ldots, x_n)$.

Representation

For a fault tree and a combinational circuit standard components, called gates can be used. E.g. AND, OR, NOT are such gates.

10. Coherent systems: A system S represented by a structure function Φ is called coherent iff the following conditions hold:

(1) If $\underline{x} < \underline{y}$ then $\Phi(\underline{x}) \leq \Phi(\underline{y})$ where $\underline{x} < \underline{y}$ means $x_i \leq y_i$ for every i, and $x_i < y_i$ for at least one i.

(2) $\Phi(\underline{1}) = 1$ and $\Phi(\underline{0}) = 0$.

Note: An informal rephrasing of (1), (2) is:

(1) If a system S is functioning no transition of a component from a failed state to functioning can cause a system failure.

(2) If all components of S are failed (functioning) the system is failed (functioning).

Example: $\Phi(\underline{x}) = x_1\bar{x}_2 + \bar{x}_1x_2$ is not coherent.

11. Minimal cuts: In a coherent system all prime implicants p_j can be represented with uncomplemented variables and are called minimal cuts. (Similarly, all prime implicates can be represented with umcomplemented variables and are called minimal paths). Let $M = \{K_1, K_2, \ldots, K_l\}$ be a set of components of a coherent system S. A subset $\mathscr{C}$ of S such that S is failed if all components K_i belonging to $\mathscr{C}$ are failed is called a cut. A cut is minimal if it has no proper subsets which are also cuts. It is called minimal cut $\mathscr{C}_j$.

12. Representation of coherent systems: Every irredundant sop representation of a structure function is a union of prime implicants. If the structure function is coherent, the representation by prime implicants greatly simplifies. We quote a theorem which leads to this simplification.

Theorem: A coherent structure function $\Phi(\underline{x})$ can be represented as a sop

$$\Phi(\underline{x}) = \sum_{j=1}^{l} p_j \tag{1.3-1}$$

of prime implicants, where this representation is unique and can be written using the concept of min cuts

$$\Phi(\underline{x}) = \sum_{j=1}^{l} \prod_{K_i \varepsilon \mathscr{C}_j} x_i \tag{1.3-2}$$

where $K_i \varepsilon \mathscr{C}_j$ are the components belonging to $\mathscr{C}_j$, x_i the Boolean variables describing the states (functioning, failed) of the components.

Note, that there is only one (minimal) cover, and there are only 'essential' prime implicants which may not be replaced by any other prime implicants.

Remark: The concept of coherence may be generalized to cases where mor than two states are possible. Even then the coherent structure functions give a considerable simplification as has been shown in /5/.

1.4 Search for Min Cuts

There are various approaches to find all min cuts. It will be sufficient to describe one algorithm in detail /6/.

Top-down-algorithm (Fussel's Algorithm)

Assume a fault tree which is given by A_o (vertex without successor), A_i (vertices with successors and predecessors, gates), x_k (vertices without predecessors, independent variables). Note: The programming contains further details which are not shown here.

Step 0 Start at top A_o.

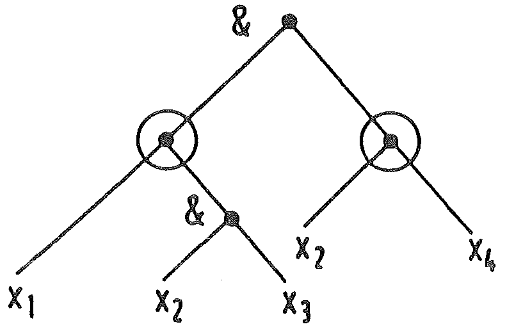
Step 1 Search for predecessors of $A_i (i=1,2,...)$
Define predecessors of A_i: $(A_i^1, A_i^2) = \text{pred } A_i$.

Step 2 If A_i is an OR gate, we get $A_i^1 + A_i^2 = A_i$,
If A_i is an AND gate, we get $A_i^1 \cdot A_i^2 = A_i$.
Rename A_i^1, A_i^2 .

Step 3 Multiply out all identified terms to obtain a sum of products. If the sum of products contains still gates (A_i) goto 1, else goto 4.

Step 4 Simplify the sum of products expression, drop repeated variables, make absorptions.

Example



$\Phi(\underline{x}) = (x_1 + x_2x_3)(x_2 + x_4)$ leads to the following min cut representation

$$\sum_{j=1}^{3} p_j = x_1x_2 + x_1x_4 + x_2x_3$$

This algorithm may be improved for systems with a high number of min cuts, e.g. by taking into account subtrees which have no replicated vertices.

Bottom-up-algorithm

This algorithm is due to Bennetts /7/ and has been improved by Nakashima /8, 9/. It begins with primary events (vertices without predecessors) and works upward to the top event. This algorithm is based on the principle of discarding redundant terms from a sop form to yield a reduced form. The improved bottom up algorithm can reduce the work needed for discarding redundant terms.

1.5 Noncoherent case

If a system is noncoherent, the approach using min cuts (min paths) has to be replaced by a search for prime implicants. Many methods have been proposed, mainly in relation to switching theory. We will give one of these methods which makes no use of minterms /10/.

Nelson's algorithm

F is already available as a polynomial (sop), which is in general not yet an expression with prime implicants.

Step 1 Complement F, and obtain $\bar{F}$ (applying De Morgan's rules). Expand $\bar{F}$ into a sop

- Drop zero products ($x\bar{x} = 0$), repeated literals ($xx = x$), make absorptions ($x + xy = x$). The result is $\bar{\Phi}$.

Step 2 Complement $\bar{\Phi}$, and obtain ϕ (applying De Morgan's rules). Expand ϕ into a sop.

- Drop zero products, repeated literals, make absorptions.

The result is Σp_i, the sum of all prime implicants, and only of prime implicants.

Example: F is available as polynomial.

Step 1 $F = x_1x_2 + \bar{x}_2x_3x_4 + x_3\bar{x}_4$

Complement F and obtain $\bar{F} = (\bar{x}_1 + \bar{x}_2)(x_2 + \bar{x}_3 + \bar{x}_4)(\bar{x}_3 + x_4)$

Expand and simplify: $\bar{\Phi} = \bar{x}_1x_2x_4 + \bar{x}_1\bar{x}_3 + \bar{x}_2\bar{x}_3$

Step 2 Complement $\bar{\Phi}$ and obtain $\Phi = (x_1 + \bar{x}_2 + \bar{x}_4)(x_1 + x_3)(x_2 + x_3)$

Expand and simplify: $\Sigma p_i = x_1x_2 + x_1x_3 + \bar{x}_2x_3 + x_3\bar{x}_4$

This algorithm can be improved in various ways, e.g. by factoring the Boolean expressions during the two steps.

2. Probabilistic Evaluation

2.1 Basic Concepts and Notations

We describe the behavior of a component which can be in a finite number of states, preferably in two states: up (functioning) or down (failed).

We describe the states by indicator variables. There is a one-one-relation between indicator variables and Boolean variables (see e.g. Barlow /2/).

Thus we get for an indicator variable $x_i^!(t)$ the following realizations:

$$x_i^!(t) = \genfrac{}{}{0pt}{}{1}{0} \text{ if component } i \text{ is } \genfrac{}{}{0pt}{}{\text{down}}{\text{up}} \text{ at time } t \qquad (2.1\text{-}1)$$

We describe the behavior of a repairable component by an alternating renewal process. Later on, it will be shown, how a system, given by a structure function, can also be represented using alternating renewal processes for components.

Availability and Reliability

We introduce a few basic quantities for reliability.

Life time distribution

Assume a component which may be modeled by a life time distribution F(t):

$$F(t) = P\,\{T \leq t\}, \qquad (2.1\text{-}2)$$

where the r.v. T is the component's life time.

Reliability

We introduce the reliability of a component R(t) as

$$R(t) = 1 - F(t) \qquad (2.1\text{-}3)$$

Note:

- For $t = o$, a component is up with probability 1.
- For $t = \infty$, a component is down with probability 1.

It is sometimes convenient to use an interval reliability (see sect. 3.3).

Availability

We introduce the availability of a component A(t)

$$A(t) = P\,\{x'(t) = 0\}, \qquad (2,1\text{-}4)$$

i.e. the probability that a component is up at time t.

Unavailability

$$\bar{A}(t) = P\,\{x'(t) = 1\}, \qquad (2.1\text{-}5)$$

i.e. the probability that a component is down at time t. Clearly

$$A(t) + \bar{A}(t) = 1$$

To obtain non-trivial statements on availability and other quantities related a few concepts of renewal theory are required.

2.2 Renewal Processes

Renewal theory deals with independent identically distributed (i.i.d.) random variables, and with the number of renewals /2/, /11/, /12/.

Assume a sequence of r.v. T_o, T_1, T_2, ... which may be represented as life times. Upon failure replacement is done in a negligible time. Let N(t) be the number of renewals in the interval (0,t) and let

$$S_n = \sum_{i=o}^{n} T_i \qquad (2.2\text{-}1)$$

Definition

Let T_o, T_1, T_2, ... be nonnegative independent r.v. with a cumulative distribution function

for T_o : $F_A(t)$

for T_i : $F(t)$ $(i \geq 1)$.

Then the sequence of the r.v. $\{T_i; i \geq 0\}$ (or equivalently $\{S_i; i \geq 0\}$) defines a renewal process.

Note: The process $\{N(t); t \geq 0\}$ is known as renewal counting process. Its relation to a renewal process is due to the equivalence:

$$\{N(t) = n\} \quad \text{iff} \quad \{S_n \leq t < S_{n+1}\}. \tag{2.2-2}$$

A few concepts related to renewal processes

1. Ordinary renewal process: If $F_A(t) = F(t)$, the process will be called ordinary.
2. Stationary renewal process: If the relation

$$F_A(t) = \frac{1}{u} \int_o^t (1-F(x))dx \tag{2.2-3}$$

with $\qquad u = E(T_i) < \infty \quad$ (for i = 1,2,3,..)

holds, the process is called stationary.

3. Renewal function: The expected number of renewals in the interval (0,t),

$$H(t) = E(N(t)) \tag{2.2-4}$$

is called renewal function. Note that

$$H(t) = \sum_{k=1}^{\infty} k \cdot P\{N(t) = k\} \tag{2.2-5}$$

If H(t) has a derivative,

$$h(t) = \frac{dH(t)}{dt} \tag{2.2-6}$$

is called renewal density. It always exists for our purposes.

4. Evaluation of a renewal function: H(t) may be defined by an integral equation of renewal type or by an infinite series of convolutions which are needed for $P\{N(t) \geq k\}$. It is convenient to evaluate H(t) in the Laplace domain. For ordinary renewal processes we get

$$H^*(s) = \frac{F^*(s)}{1-sF^*(s)} \tag{2.2-7}$$

where * refers to the Laplace transform.

Poisson process

For a sequence, where all T_i are i.i.d. with $F(t) = 1-e^{-\lambda t}$ we get a Poisson process, where λ is a fixed parameter. For the renewal function we get

$$H(t) = \lambda t. \tag{2.2-8}$$

It is also interesting to note the relation to other definitions of the Poisson process /2/. For the renewal counting process related to the Poisson process we note: $\{N(t) - N(t_o); t \geq t_o\}$ constitutes a Markow process.

We also note:

$$M_t = N_t - N_{t_o} - \lambda t \tag{2,2-9}$$

is both a Markow process and a Martingale /13/.

Superposition of processes

Assume there are m independent components which fail at random times, where the failed components are repaired instantaneously. Assume that each single component generates a renewal process. The failures of all m components may be again modeled by a stochastic process (a point process) which is in general no longer a renewal process. However, we may note for the superposition of Poisson processes:

If there are m Poisson processes (with mean values u_i ($i = 1,2,\dots,m$)), then the renewal function for the superposition of m Poisson processes $H_s(t)$ (expected number of all renewals in $(0,t)$) may be given:

$$H_s(t) \quad = \quad \sum_{i=1}^{m} H_i(t) \qquad (2.2\text{-}11)$$

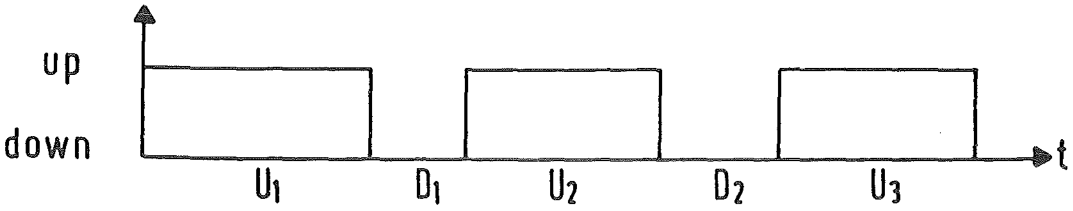
where $H_i(t)$ is the renewal function for the i^{th} Poisson process, with

$$H_i(t) \quad = \quad \frac{t}{u_i} \qquad (2.2\text{-}12)$$

This is relevant for a series system.

2.3 Alternating Renewal Processes

We consider a component which can be in one of two states, up and down, but is no longer repaired instantaneously /2/, /11/, /12/. Thus we have this realization:



Initially it is up and remains for a time U_1, then it goes down and remains down for a time D_1 etc.

The time intervals

$$T_i \equiv (U_i + D_i) \qquad i = 1,2,3,\dots \qquad (2.3.\text{-}1)$$

are assumed to be mutually independent.

Let U_i($i = 1,2,3,\dots$) be distributed with $F(t)$,

let D_i($i = 1,2,3,\dots$) be distributed with $G(t)$, and

let $T_i \equiv (U_i + D_i)$ ($i = 1,2,3,\dots$) be distributed with $F_T(t)$ ($i = 1,2,3,\dots$).

Then the sequence of r.v. $\{T_i;\ i \geq 1\}$ defines an alternating renewal process, where

$$F_T(t) = P\,\{T_i \leq t\} \;=\; \int_o^t f(x)\ G(t-x)dx. \qquad (2.3\text{-}2)$$

A few concepts related to alternating renewal processes

1. Ordinary renewal process: The definition already refers to the ordinary process.

2. Mean values (u, d):

$$\begin{aligned} &(a) \quad u = E(U_i) \\ &(b) \quad d = E(D_i) \qquad (i = 1,2,\ldots) \qquad (2.3\text{-}3) \\ &(c) \quad u+d = E(T_i) \end{aligned}$$

3. Renewal function: We get for the mean number of failures H(t) (assuming an up state for t=0):

$$H^*(s) = \frac{F^*(s)}{1-f^*(s)\ g^*(s)} \qquad (2.3\text{-}4)$$

Relation to Point Processes

It is interesting to note that the abovementioned renewal processes are special cases of point processes. A point process over the half line $[0,\infty)$ can be viewed as follows:

(a) as a sequence of nonnegative r.v.: $T_o, T_1, T_2, \ldots.$

(b) as an associated counting process N_t where

$$N_t = \begin{array}{lll} n & \text{if} & t\varepsilon[T_n, T_{n+1}) \\ \infty & \text{if} & t = \lim T_n = \infty \end{array} \qquad (2.3\text{-}5)$$

see also (2.2-2) (renewal counting process). The Poisson process is a well known example for a point process. (2.2-9) which relates the counting process N_t, the intensity λ and the martingale M_t is very useful (see Brêmaud /24/ and section 4).

Availability of a Component

We now obtain a few relations of Availability and alternating renewal processes. Assume a component which is in an up state for t=0. The time U_1 to the first failure is distributed as $F_A(t) = 1-\bar{F}_A(t)$. The times U_i (i>1) (referring to operation) are distributed as $F(t) = 1-\bar{F}(t)$ and the times D_i are all distributed as G(t) (see Fig. 3 and (2.3-1)). Then we obtain for the availability A(t) the following formulas:

$$A(t) = \bar{F}_A(t) + \int_o^t \bar{F}(t-x)\,dH(x) \qquad (2.3\text{-}6)$$

Example: For an alternating renewal process where up and down times are exponentially distributed, we get

$$A(t) = \frac{\rho}{\lambda+\rho} + \frac{\rho}{\lambda+\rho} e^{-(\lambda+\rho)t}$$

Asymptotic behavior

As applications of the key renewal theorem we get the following relations (see (2.2-3), (2.3-10):

$$\text{(a)} \lim_{t\to\infty} \frac{H(t)}{t} = \lim_{t\to\infty} h(t) = \frac{1}{u+d} \tag{2.3-7}$$

(The same holds for $\tilde{H}(t)$, $\tilde{h}(t)$).

$$\text{(b)} \lim_{t\to\infty} (H(t+x) - H(t)) = \frac{x}{u+d} \text{ for all } x > 0 \tag{2.3-8}$$

$$\text{(c)} \lim_{t\to\infty} A(t) = \frac{u}{u+d} \tag{2.3-9}$$

An interpretation of renewal function and density

For the application of renewal function and density to fault tree evaluation the following notation is convenient. It is possible to understand the expected number of failures (repairs) of a component i of a system as follows:

$$W_i^{01}(t) = E \text{ (Number of failures in } (0,t) \text{ for component i)} \tag{2.3-10}$$

$$W_i^{10}(t) = E \text{ (Number of repairs in } (0,t) \text{ for component i)} \tag{2.3-11}$$

corresponding to $H(t)$, $\tilde{H}(t)$ respectively. Moreover:

$$w_i^{01}(t)dt = P\{\text{component i fails in } (t, t+dt)\} \tag{2.3-12}$$

where $w_i^{01}(t)$ is the failure intensity,

$$w_i^{10}(t)dt = P\{\text{component i is repaired in } (t,t+dt)\} \tag{2.3-13}$$

where $w_i^{10}(t)$ is the repair intensity.

Similarly,

$w_s^{01}(t)$ is the failure intensity of the system, $w_s^{10}(t)$ the repair intensity.

Note

1. The failure intensity-notation replaces for the rest for this representation the usual $h(t)$.

2. The failure intensity may be easily generalized to a transition rate for a finite number of states /5/, /14/.

3. Assume that up times and down times are exponentially distributed. Then we get

$$w_i^{01}(t) = \lambda_i A_i(t)$$
$$w_i^{10}(t) = \mu_i \bar{A}_i(t) \qquad (2.3\text{-}14)$$

where λ_i,(μ_i) is the failure rate (repair rate) of i.

2.4 Stochastic Modeling of a System

Based on 2.1-2.3 we now introduce concepts which are useful for reliability evaluations of systems. We assume a coherent system (C,ϕ) with n components /12/.

Alternating renewal process

1. Component i is replaced at failure (not instantaneously) thus generating an alternating renewal process, where renewal densities are $w_i^{01}(t)$, $w_i^{10}(t)$ (i = 1,2,...,n).
2. For a stationary process we have ((2.3-11),(2.3-9)):

$$w_i^{01} = w_i^{10} = A_i/u_i = 1/(u_i + d_i) \qquad (2.4\text{-}1)$$

3. We assume that components i,j (i ≠ j) are statistically independent and that

$$w_i^{01}(t)\ w_j^{01}(t)\ (dt)^2 = o(dt) \quad \text{for } i \neq j \qquad (2.4\text{-}2)$$

$$w_i^{01}(t)\ w_j^{10}(t)\ (dt)^2 = o(dt) \qquad (2.4\text{-}3)$$

where o(u) is the Landau symbol, i.e. for a function f we get

$$f(u) = o(u) \quad \underline{\text{iff}} \quad \lim_{u \to 0} \frac{f(u)}{u} = 0 \qquad (2.4\text{-}4)$$

Thus it is possible to exclude that two failures or one failure and one repair occur at 'the same time'.

4. Of course, a coherent system will in general not follow a renewal process.

Unavailability

The state $X_s'(t)$ of the system can be expressed in terms of component states, $X_1'(t),\dots,X_n'(t)$:

$$X_s'(t) = \Phi(X_1'(t),\dots,X_n'(t)) \qquad (2.4\text{-}5)$$

It follows that unavailability $\bar{A}_s(t)$ of the system at time t is given as

$$\bar{A}_s(t) = E(X_s'(t)) = h(\bar{A}_1(t),\dots,\bar{A}_n(t)) \qquad (2.4\text{-}6)$$

where h is the 'reliability function' of system (C,Φ), i.e. the (point-) unavailability at time t /2/, /15/.

Limiting unavailability

Let U_{ji} represent the i th up time for component j with distribution F_j (mean u_j), and D_{ji} represent the i th down time for component j with distribution G_j (mean d_j), for j = 1,2,...,n, i = 1,2,3,... .

Since h is multilinear in its arguments, the stationary unavailability $\bar{A}_s$ is, for nonlattice distributions of F_j, G_j,

$$\bar{A}_s = h(\frac{d_1}{u_1+d_1},\dots,\frac{d_n}{u_n+d_n}) \qquad (2.4\text{-}7)$$

For AND and OR-gates we get as unavailability:

1. AND-gate

$$\bar{A}_s(t) = P\{X_1'(t)\cdot X_2'(t) = 1\} = \bar{A}_1(t)\ \bar{A}_2(t) \qquad (2.4\text{-}8)$$

$$\bar{A}_s = d_1/(u_1+d_1)\cdot d_2/(u_2+d_2) \qquad (2.4\text{-}9)$$

2. OR-gate

$$\bar{A}_s(t) = P\{1-(1-X_1'(t))(1-X_2'(t)) = 1\} = 1-(1-\bar{A}_1(t))(1-\bar{A}_2(t)) \qquad (2.4\text{-}10)$$

$$\bar{A}_s = 1 - u_1/(u_1+d_1)\cdot u_2/(u_2+d_2) \qquad (2.4\text{-}11)$$

Failure intensity

The evaluation of failure intensity of a system is related to assumptions (2.4-1)-(2.4-3) and to the concept of a critical component.

Critical component

A coherent system is in a state where component j is critical iff for the structure function Φ

$$\Phi(1_j,\underline{x}) - \Phi(0_j,\underline{x}) = 1 \qquad (2.4\text{-}12)$$

holds, where $(1_j,\underline{x}) = (x_1,x_2,\dots,x_{j-1},1,x_{j+1},\dots,x_n)$, similarly $(0_j,\underline{x})$. The system fails, if component j fails. The state of the system is adjacent to system failure. The probability, that a system is in a state where component j is critical, may be given as

$$I_j = P\{\Phi(1_j,\underline{x}) - \Phi(0_j,\underline{x}) = 1\} \qquad (2.4\text{-}13)$$

We get with the reliability function $h(\underline{p})$,

$$h(\underline{p}) = p_j\ h(1_j,\underline{p}) - (1-p_j)\ h(0_j,\underline{p}) \qquad (2.4\text{-}14)$$

$$I_j = \frac{\partial\ h(\underline{p})}{\partial\ p_j} = h(1_j,\underline{p}) - h(0_j,\underline{p}) \qquad (2.4\text{-}15)$$

This is also known as Birnbaum's importance measure which may be used for sensitivity analysis. But here it is of central relevance for evaluation of our fault trees.

A fundamental relation

The following theorem shows a fundamental relation between the failure intensity of a system and its components /15/.

Theorem

If a system is coherent, we get

$$w_s^{01}(t) = \sum_{i=1}^{n_c} I_i(t)\, w_i^{01}(t) \tag{2.4-16}$$

where $I_i(t) = \dfrac{\partial h(\bar{\underline{A}}(t))}{\partial \bar{A}_i(t)}$,

and the summation has to be taken over all states i ($1 \leq i \leq n_c$) in which the failure of a component is critical.

Proof: Since $I_i(t)$ may be represented as the probability that the system is in a state where component i is critical, the probability that a system failure in (t,t+dt) is caused by a failure of component i, is given as $I_i(t)w_i^{01}(t)dt$ where $w_i^{01}(1)$ = P{component i fails in (t,t+dt)} (2.3-27). The simultaneous occurrence of two component failures may be regarded as small compared to $w_i^{01}(t)dt$ (2.4-2). Thus the probability for any system failure in (t,t+dt) is

$$w_s^{01}(t)dt = \sum_{i=1}^{n_c} I_i(t)\, w_i^{01}(t)dt$$

Note: It is important to note that only $w_i^{01}(t)$ (rather than $w_i^{10}(t)$) will be needed for coherent systems. For the noncoherent case, we will also have a dependence on $w_i^{10}(t)$. This can be generalized to the multivalued case (see Barlow /5/, Murchland /14/).

Examples

1. AND-gate

Note, that for an AND-gate components i = 1,2 are critical. They are also predecessors of this gate.

$$w_s^{01}(t) = \bar{A}_2(t)\, w_1^{01}(t) + \bar{A}_1(t)\, w_2^{01}(t) \tag{2.4-17}$$

2. OR-gate

Note, that for an OR-gate components i = 1,2 are critical.

$$w_s^{01}(t) = (1 - \bar{A}_2(t))\, w_1^{01}(t) + (1 - \bar{A}_1(t))\, w_2^{01}(t) \tag{2.4-18}$$

3.

For a fault tree without replications the two abovementioned relations are sufficient to evaluate $w_s^{01}(t)$ in terms of all predecessors. Only a recursive procedure, applying the theorem for all gates is needed. However, for trees with replications, we need further considerations.

3. Evaluation with Min Cuts and Min Paths

3.1 Basic Concepts

Consider a coherent system which can be represented using, min cuts $\mathcal{C}_j$ or min paths $\mathcal{P}_k$. We denote by x_i' an indicator variable (see (2.1-1) and use the notations:

Product: $$\prod_{i=1}^{n} x_i' \qquad (3.1\text{-}1)$$

Coproduct: $$\coprod_{i=1}^{n} x_i' = 1 - \prod_{i=1} (1-x_i') \qquad (3.1\text{-}2)$$

(3.1-1) and (3.1-2) is related to Boolean products and Boolean sum respectively. For the reliability function $h(\underline{p})$ (2.4-7) we may write:

$$E(\prod_{k=1}^{m} \coprod_{i\varepsilon\mathcal{P}_k} x_i') = h(\underline{p}) = E\,(\coprod_{j=1}^{l} \prod_{i\varepsilon\mathcal{C}_j} x_i') \qquad (3.1\text{-}3)$$

where $\mathcal{P}_k$, ($\mathcal{C}_j$) refers to min paths (min cuts).

Note that this is related to two major forms for a Boolean expressions: The sum of products form (r.h.s.) and the product of sums form (l.h.s.) which are equivalent. If there is a coherent structure, we get in general the following bounds

$$\prod_{k=1}^{m} \coprod_{i\varepsilon\mathcal{P}_k} p_i \le h(\underline{p}) \le \coprod_{j=1}^{l} \prod_{i\varepsilon\mathcal{C}_j} p_i \qquad (3.1\text{-}4)$$

However, for noncoherent structures, the bounds will not hold in general /2/.

The time to failure for a Coherent System

Let t_i be the time to failure of the i-th component (i=1,2...,n), and $\tau_\phi(t)$ the time to failure of a coherent system (C,ϕ) with structure function ϕ.

We give now a result which is related to (3.1-3) but not based on Boolean variables.

Theorem: If (C,ϕ) is a coherent system with minimal paths $\mathcal{P}_k$ (k=1,2,...,m) and minimal cuts $\mathcal{C}_j$ (j=1,2,...,l).

Then

$$\max_{1\le k\le m}\ \min_{i\varepsilon\mathcal{P}_k} t_i = \tau_\phi(t) = \min_{1\le j\le l}\ \max_{i\varepsilon\mathcal{C}_j} t_i \qquad (3.1\text{-}5)$$

Proof: A coherent system fails when the first minimal cut $\mathcal{C}_j$ fails . A parallel structure fails when the last component i of this cut $\mathcal{C}_j$ fails. (A similar argument holds for minimal paths).

(3.1-5) is of interest for methodological considerations (see section 4.1 on systems evolution).

3.2 Inclusion - Exclusion - Principle

It is convenient to have a procedure to evaluate complex fault trees, where (3.1-3) would be impractical. In general, an exact evaluation is not feasible. But it is possible to obtain bounds for unavailability, failure intensity etc. as will be discussed in sect. 3.3 .

Now the inclusion-exclusion-principle (Poincaré's theorem) will be given.

In a discrete probability space (i.e. with countable elementary events) we get the following theorem:

Theorem: Let A_1, A_2, ..., A_n be events. Then we get

$$P\{\bigcup_{i=1}^{n} A_i\} = \sum_{i=1}^{n} P\{A_i\} - \sum_{i<j} P\{A_i\, A_j\} + \ldots + (-1)^{n-1} P\{A_1\, A_2 \ldots A_n\} \,. \qquad (3.2\text{-}1)$$

This is a theorem which applies to events contained in a discrete probability space. Then it also applies to indicator variables and to events such as 'min cut failed'. Moreover, Poincaré's theorem can be restated for expectations $E(\bigcup_{i=1}^{n} A_i)$. As a corollary we note:

Corollary: We get upper (lower) bounds for $P\{\bigcup_{i=1}^{n} A_j\}$:

$$\begin{aligned} P\{\bigcup_{i=1}^{n} A_j\} &\leq \sum_{i=1}^{n} P\{A_i\} \\ P\,\{\bigcup_{i=1}^{n} A_j\} &\geq \sum_{i=1}^{n} P\{A_i\} - \sum_{i<j} P\{A_i\, A_j\} \end{aligned} \qquad (3.2\text{-}2)$$

Relation to combinatorics

The inclusion-exclusion principle is related to a fundamental enumeration procedure. This can be shown by the following relation:

Theorem: Let a_1, a_2, ..., a_n be real numbers. Then

$$(1-a_1)(1-a_2)\ldots(1-a_n) = 1 - \sum_{i=1}^{n} a_i + \sum_{i<j} a_i\, a_j - \ldots + (-1)^n a_1\, a_2 \ldots a_n \qquad (3.2\text{-}3)$$

Proof: Induction. This theorem illustrates the relation between co-product and inclusion exclusion.

3.3 Evaluation with Bounds

Usually, the exact formula of inclusion exclusion needs a large amount of computation. Therefore, bounds are required. This will be demonstrated for a fault tree represented by min cuts, where all components are repairable. The usefulness of bounds and/or approximations will be discussed.

For evaluation the following steps are required:

Algorithm

Step 1 Search for min cuts (by top down or bottom up algorithm)

Step 2 Bring the Boolean polynomial (min cuts) into a disjoint form (using the inclusion exclusion principle).

Step 3 Evaluate unavailability and failure intensity as a function of life and repair distribution (e.g. with mean uptime u, mean downtime d)

For a detailed presentation see Nakashima /8/ and Olmos, Wolf /18/.

Simple systems

Parallel system

Let C be the event that 'parallel system $\mathcal{C}$ is down'. $\mathcal{C}$ has n components and is represented by an AND-gate (see 3.1-4):

$$\bar{A}(t) = P\{C\} = \prod_{i\varepsilon\mathcal{C}} \bar{A}_i(t) \qquad (3.3\text{-}1)$$

As application of theorem (2.4-16) we get:

$$w^{01}(t) = \bar{A}(t) \sum_{i=1}^{n} \frac{w_i^{01}(t)}{\bar{A}_i(t)} \qquad (3.3\text{-}2)$$

For the stationary state, we get

$$w^{01} = \bar{A} \sum_{i=1}^{n} \frac{1}{d_i} \qquad (3.3\text{-}3)$$

where $w_i^{01} = 1/(u_i + d_i) = \bar{A}_i/d_i$.

Series system

Let T be the event that 'series system $\mathcal{T}$ is down'. $\mathcal{T}$ is represented by an OR-gate.

$$\bar{A}(t) = P\{T\} = 1 - \prod_{i\varepsilon\mathcal{T}}(1-\bar{A}_i(t)) \qquad (3.3\text{-}4)$$

$$w^{01}(t) = (1-\bar{A}(t)) \sum_{i=1}^{n} \frac{w^{01}(t)}{1-\bar{A}_i(t)} \qquad (3.3\text{-}5)$$

For the stationary state we get

$$w^{01} = (1-\bar{A}) \sum_{i=1}^{n} \frac{1}{u_i} \qquad (3.3\text{-}6)$$

where $w_i^{01} = 1/(u_i+d_i) = A_i/u_i$.

Note also relation to simple trees. We obtain the following relations as a consequence of theorem (3.2-1) and (3.2-2).

Theorem

We assume a coherent fault tree with min cut representation. Let the K_i be independent and let C_j be the event 'the min cut $\mathcal{C}_j$ fails ' $(j = 1,\dots,m)$. Upper bounds and lower bounds for unavailability $\bar{A}_s(t)$ are:

$$\text{(a)}\quad \bar{A}_s(t) \leq \sum_{j=1}^{m} P\{C_j\} \tag{3.3-7}$$

$$\text{(b)}\quad \bar{A}_s(t) \geq \sum_{j=1}^{m} P\{C_j\} - \sum_{j=1}^{m-1} \sum_{k=j+1}^{m} P\{C_jC_k\} \tag{3.3-8}$$

where C_jC_j is the event 'intersection of $\mathcal{C}_j$ and $\mathcal{C}_k$ fails' (where all replicated variables occur only once).

Note: For the r.h.s. of (3.3-8) the maximum difference from the exact value for $\bar{A}_s(t)$ is $\sum_{j=1}^{m-1} \sum_{k=j+1}^{m} P\{C_jC_k\}$, provided we have a coherent system.

Theorem

Let K_i be independent and in a stationary state. Then we get for w_s^{01} these relations: Upper bounds and lower bounds for failure intensity w_s^{01} are:

$$\text{(a)}\quad w_s^{01} \leq \sum_{j=1}^{m} P\{C_j\} \sum_{i\in\mathcal{C}_j} \frac{1}{d_i} \tag{3.3-9}$$

$$\text{(b)}\quad w_s^{01} \geq \sum_{j=1}^{m} P\{C_j\} \sum_{i\in\mathcal{C}_j} \frac{1}{d_i} - \sum_{j=1}^{m-1} \sum_{k=j+1}^{m} P\{C_jC_k\} \sum_{i\in\mathcal{C}_j \mathcal{C}_k} \frac{1}{d_i} \tag{3.3-10}$$

where $w_s^{01} = 1/(u_i+d_i) = \bar{A}_i/d_i$ for all i.

Note: For the r.h.s. of (3.3-9) the maximum difference from the exact for w_s^{01} is given by the second term of the r.h.s. of (3.3-10).

Expected number of failures

In the stationary state, we get for the expected number of failures in the interval $(0,t)$

$$W_s^{01}(t) = \int_0^t w_s^{01} dt' = t \cdot w_s^{01} \tag{3.3-11}$$

This is

$$W_s^{01}(t) \leq t \sum_{j=1}^{m} P\{C_j\} \sum_{i\in\mathcal{C}_j} \frac{1}{d_i} \tag{3.3-12}$$

Unreliability

For the unreliability of a repairable system we need a few events to relate it to other concepts already introduced:

$S(t) = \{\text{the system is up at time } t\}$

$N_c(t,\ t+\tau) = \{\text{no system transition from up to down in } (t,\ t+\tau) \mid \text{the system is up at } t\}$

$N_i(t,\ t+\tau) = \{i \text{ system transitions from up to down in } (t,\ t+\tau)\}$

We define the unreliability $\bar{R}(t,\ t+\tau)$:

$$\bar{R}_s(t,\ t+\tau) = 1 - P\{N_o(t,\ t+\tau) \mid S(t)\} = \bar{P}\{N_c(t,\ t+\tau)\} \tag{3.3-13}$$

i.e. the probability that there are more than zero transitions from up to down in $(t,\ t+\tau)$ conditional on the system being up at t. Note that this differs from the usual definition of unreliability. There exists no analytical method for calculating the unreliability for general coherent systems with repairable components /19/, /15/. However, using W_s^{01} and A_s a bound may be given.

Theorem

For system unreliability $\bar{R}_s(t,\ t+\tau)$, conditional on the system being up at time t, a bound is:

$$\bar{R}_s(t,\ t+\tau) \leq W_s^{01}(t,\ t+\tau)\ /\ A_s(t) \tag{3.3-14}$$

Proof: Due to (3.3-13) we obtain

$$\bar{R}_s(t,\ t+\tau) = 1 - P\{N_o(t,\ t+\tau) | S(t)\} = \sum_{i=1}^{\infty} P\{N_i(t,\ t+\tau) | S(t)\} \leq \sum_{i=1}^{\infty} iP\{N_i(t,\ t+\tau) | S(t)\}$$

Next we add to the r.h.s. the expected number of failures, conditional on $\bar{S}(t)$. By a suitable multiplication we get, using the total law of probability

$$\sum_{i=1}^{\infty} iP\{N_i(t,\ t+\tau) | S(t)\} \leq \frac{1}{A(t)}\left[\sum_{i=1}^{\infty} iP\{N_i(t,\ t+\tau) | S(t)\}A(t) + \Sigma iP\{N_i(t,\ t+\tau) | \bar{S}(t)\}\bar{A}(t)\}\right]$$

$$= \frac{1}{A(t)} \sum_{i=1}^{\infty} iP\{N_i(t,\ t+\tau)\} = \frac{W_s^{01}(t,t+\tau)}{A_s(t)} \qquad \text{(see (2.2-5))}.$$

An interesting special case is this:

$$\bar{R}_s(\tau) = \bar{R}_s(0,\tau) \leq W_s^{01}(\tau), \tag{3.3-15}$$

when all components are intact at t=0. We give an application of this theorem: For a parallel system with n components (j=1, ..., n) we get ((3.3-3), (3.3-14) for the stationary state:

$$\bar{R}_s(t,\ t+\tau) \leq \tau \prod_{j=1}^{n} \frac{d_j}{u_j} \left(\sum_{i=1}^{n} \frac{1}{d_i}\right) \tag{3.3-16}$$

where u_j mean time between failures,
d_j mean time to repair.

Note:

1. If τ is large compared to $\max_{i=1,\dots,u} d_i$, system availability is high and the unreliability to be calculated is rather accurate.
2. A similar formula holds for the more general case of a coherent system.

A few limitations to fault tree analysis

In relation to system reliability a few remarks on the limits of fault tree analysis are in order. They have been observed by various authors /19/, /15/, /20/.

There has been a long debate on the applicability of Kinetic Tree Theory which is due to Vesely /21/. It is claimed that kinetic tree theory can evaluate system reliability by analytical means where

$$P\{\text{no system failure in } (0,t)\} = \exp(-\int_o^t \Lambda_o(x)dx) \qquad (3.3\text{-}19)$$

where

$$\Lambda_o(x)dx = P\,\{\text{system fails in } (x,\ x+dx) \mid \text{it was up at } x\} \qquad (3.3\text{-}20)$$

It has been shown by verious authors /19/, /15/ /20/ that (3.3-17) is only correct iff

$$\Lambda_o(t) = \Lambda(t) \qquad (3.3\text{-}21)$$

where

$$\Lambda(t)dt = P\,\{\text{system fails in } (t,\ t+dt) \mid \text{it never failed before } t\}. \qquad (3.3\text{-}22)$$

It can be shown that this condition is not valid in general. E.g.

1. If components are nonrepairable, (3.3-19) holds.
2. If components are repairable and in series, (3.3-19) also holds.
3. It can be shown that for a parallel system of 2 components, where the life times and the repair times are i.i.d. and exponential, two basically different results are obtained:

a) If we evaluate reliabilities on the basis of Veselys formalism or
b) if we evaluate reliabilities on the basis of a Markow process.

This counterexample (due to /15/) demonstrates that this method does not hold in general.

It can be shown that for the aysmptotic case this difference vanishes. Moreover, we get for reasonable values of t a good approximation for reliability. /20/ has discussed in detail the assumptions required for this evaluation.

Much more serious limitations for fault tree analysis arise if events are no longer statistically independent. To discuss the available methods would be beyond the scope of this lecture. See Barlow /2/.

Increasing failure rate

If materials, components or subsystems wear out with time, the class of distribution (survival functions) where the failure rate is increasing (IFR) is evidently of special interest. Ignoring the possibility of "infant mortality" this is usually a strong and natural assumption.

A component with life time distribution F(t) (2.1-2) has the IFR-property if

$$\lambda(t) = \frac{f(t)}{1-F(t)}, \quad t \geq 0 \qquad (3.3\text{-}16)$$

is increasing. A more general concept is this: A component has increasing failure rate average (IFRA) if

$$\frac{1}{t}\,\Lambda(t) = \frac{1}{t}\int_0^t \lambda(u)\,du = -\frac{1}{t}\ln\,(1-F(t)) \qquad (3.3\text{-}17)$$

is increasing.

We are not considering here the closely related DFR-concept. Birnbaum /25/ and Barlow /2/ discussed many properties of IFR, DFR and IFRA, DFRA - distributions.

The IFR - property may be related to convexity:

Theorem: A life time distribution (F(t)) is IFR iff the cumulative failure rate

$$\Lambda(t) = -\log \bar{F}(t) \qquad (3.3\text{-}18)$$

is convex in the interval where it is defined.

We recall that a convex function is

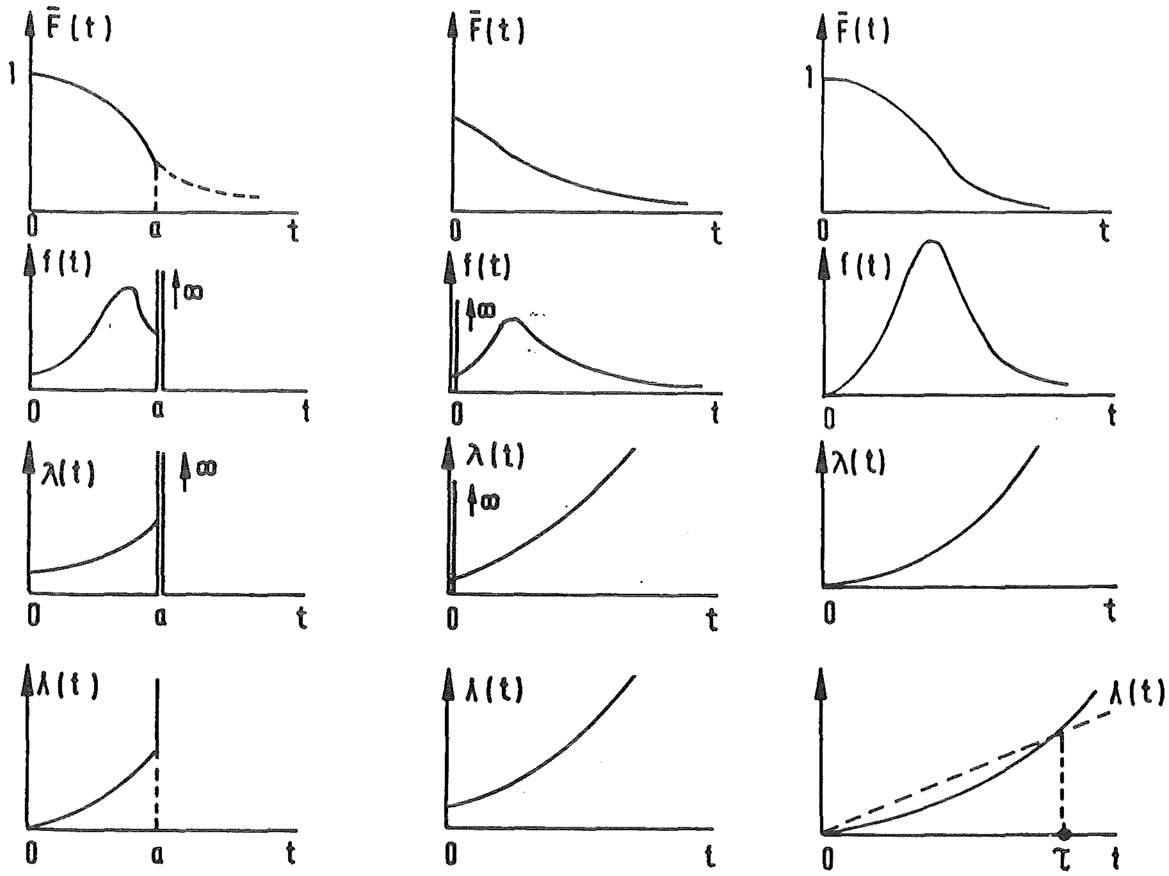
- necessary continuous and has
- at every point a left- and right-derivative which are nondecreasing.

Let us make a more general statement related to (3.3-18).

We state the following theorem:

A life time distribution (F(t)) is IFRA iff the cumulative failure rate Λ(t) is convex and passes through the origin.

A few examples will illustrate these criteria.



(a) (b) (c)

Note:

(a) Λ(t) is convex in [0,a) : IFR

(b) Λ(t) is convex but makes a jump at 0 : not IFR

(c) Λ(t) is convex and passes through the origin: IFR and IFRA.

Closure properties

We assume components with exponential survival functions (constant failure rates λ_1,λ_2).

For a series system we obtain:

$$\bar{F}(t) = e^{-\lambda_1 t} \cdot e^{-\lambda_2 t} = e^{-(\lambda_1+\lambda_2)t} \qquad (3.3\text{-}19)$$

$$\lambda_s = \lambda_1+\lambda_2 .$$

Thus λ_s is also constant.

However, for a parallel system ($\lambda_1 \neq \lambda_2$) we obtain:

$$\bar{F}(t) = 1 - (1-e^{-\lambda_1 t})(1-e^{-\lambda_2 t}) \qquad (3.3\text{-}20)$$

$$\lambda_p(t) = \frac{\lambda_1 e^{-\lambda_1 t}+\lambda_2 e^{-\lambda_2}-(\lambda_1+\lambda_2)e^{-(\lambda_1+\lambda_2)t}}{e^{-\lambda_1 t}+e^{-\lambda_2 t}-e^{-(\lambda_1+\lambda_2)t}} \qquad (3.3\text{-}21)$$

Here, $\lambda_p(t)$ is in general not constant. It is not IFR either.

But for the IFRA-property of survival function we may state:

Theorem: A system with a coherent structure function having components with IFRA-survival-functions has itself a IFRA survival function.

Example: $\lambda_p(t)$ of (3.3-21) is IFRA.

It can be seen in sect. 4.1 that the IFRA property is also related to point processes.

4. Evolution of a Coherent System

The question is as follows: If a coherent system has alternating renewal processes at the component level, what can be said regarding the evolution of the system? Evidently the following holds:

- Unavailability and failure intensity can be evaluated.
- The alternating renewal processes are a special type of point processes and of Markov renewal processes (see also sect. 2.3).

Three approaches are possible: use of stopping times, distributions of phase type, Markov renewal processes.

4.1 Times to failure and stopping times

The basic idea is very simple: Consider two components in series with random life times T_1, T_2. Then this system fails at the time

$$T_* = \inf(T_1,T_2). \tag{4.1-1}$$

Consider also two components in parallel with random life times T_1,T_2. Then this system fails at the time

$$T^* = \sup(T_1,T_2) \tag{4.1-2}$$

It is evidently possible to use the relations (4.1-1), (4.1-2) to obtain statements for a coherent system (C,ϕ).

As a stochastic concept, the stopping time T is required. The stopping time is based on the understanding that at time $t \geq 0$ it is known whether an event (component failure) occurred or not.

Def. Let F_t be a collection of events, representing the known information at time t. (F_t is also called a σ-field of events). F_t is typically the collection of events generated by one or more stochastic processes up to time t. Now let $\{F_t, t \geq 0\}$ be a family of such information collections. We shall always assume that $\{F_t\}$ is increasing, i.e. that no forgetting is allowed:

$$s \leq t \quad \Rightarrow \quad F_s \subseteq F_t \tag{4.1-3}$$

here $\{F_t,\ t \geq 0\}$ is called a history.

With this concept we may define the stopping time.

Def. Let $\{F_t\}$ be a history and T be a possible random variable. Then T is called a F_t-stopping time iff the event $\{T \leq t\}$ can be characterized by

$$\{T \leq t\} \subset F_t\ , \quad t \geq 0 \tag{4.1-4}$$

i.e. it is known at time t whether or not T has occurred.

A few properties of stopping times

A process X_t is called "adapted" to $\{F_t,\ t \geq 0\}$ if for every t, X_t is completely determined by F_t.

1. Theorem: Let X_t be a right-continuous $\mathbb{R}$-valued process adapted to F_t, and c a given real number. Define T as follows:

$$T = \begin{cases} \inf\{t | X_t \geq c\} \\ +\infty \text{ if this set is empty} \end{cases} \qquad (4.1\text{-}5)$$

Then T is a F_t - stopping time.

Proof: See Bremaud /24/. We note that T under conditions (4.1-5) is a "first passage time". An important special case of a first passage time will be used in sect. 4.2 (phase type distributions).

2. Relation of two stopping times: If T_1, T_2 are F_t - stopping times, then

$$\begin{aligned} T_1 \wedge T_2 &:= \inf\,(T_1,T_2) \\ T_1 \vee T_2 &:= \sup\,(T_1,T_2) \end{aligned} \qquad (4.1\text{-}6)$$

are also F_t - stopping times.

Note that $\wedge (\vee)$ is here not the conjunction (disjunction) but a useful symbol for inf(sup), corresponding to series (parallel) systems life time. We can obtain (4.1-6) considering

$$\begin{aligned} \{T_1 \wedge T_2 \leq t\} &= \{T_1 \leq t\} \cup \{T_2 \leq t\} \\ \{T_1 \vee T_2 \leq t\} &= \{T_1 \leq t\} \cap \{T_2 \leq t\} \end{aligned} \qquad (4.1\text{-}7)$$

The relation (4.1-6) may be generalized as follows: Let T_i (i=1,2,...,m) stopping times. Then

$$T^* = \sup_{1 \leq i \leq m} T_i \qquad (4.1\text{-}8)$$

$$T_* = \inf_{1 \leq i \leq m} T_i \qquad (4.1\text{-}9)$$

is also a stopping time. (Brêmaud /24/).

Combining these relations we obtain the following theorem:

Theorem: Let (C,ϕ) be a coherent structure where $T_1,\ldots,T_n$ are stopping times (life times). Then T_ϕ,

$$T_\phi = \inf_{1 \leq j \leq l} \; \sup_{i \in \mathcal{C}_j} T_i \qquad (4.1\text{-}10)$$

is also a stopping time.

Here $\mathcal{C}_j$ is a minimal cut and 1 is the number of minimal cuts of (C,ϕ). This theorem follows from formulas (4.1-8), (4.1-9) and from equation (3.1-5). (See also Greenwood /26/).

Thus our question regarding the time to system failure has been answered. It is interesting to know also the kind of stochastic process which describes this system.

Point Processes

It is possible to define point processes as a sequence of stopping times. We can make (with (4.1-10)) the following statement.

If on the component level we have point processes (renewal processes are a special type of point processes)

then on the system level (for a coherent system (C,ϕ)) we also have a point process.

These considerations are due to Arjas /27/ and Greenwood /26/. But this is not the place to discuss this in detail.

Now let us make a few remarks how to construct such a point process.

It is possible to characterize a point process as follows:

$$M_t = N_t - \int_0^t \lambda_s \, ds \qquad (4.1\text{-}11)$$

where N_t is the counting process associated to apoint process (2.3-5),

$\int_0^t \lambda_s \, dt$ is a compensator (integral of the intensity) of a point process,

M_t is a martingale.

This is called decomposition of a point process (see eg. Doob /13/, Brémaud /24/). Example: For a Poisson-process we have:

$$M_t = N_t - \lambda t \qquad (2.2\text{-}9)$$

A few remarks on point processes

1. Based on the decomposition which has been sketched (see (4.1-11)) it is possible to do some considerations which come very close to the IFRA-properties of systems (see also closure property (3.3-21).

This has been discussed in detail by Arjas /27/ and Greenwood /26/.

We will learn more about certain closure properties of distributions in the next section.

2. Using a "marked point process" we can formalize the considerations we already mentioned referring to a coherent system (Greenwood /26/). Note that the point process approach uses classical methods such as imbedded Markow chains, imbedded Markow processes and semi-Markow processes (see König, Stoyan /28/, Cox and Miller /29/). It is sometimes perferable to use methods which are specific to point processes.

3. It has been shown by various anthors (Arndt, Franken /30/, Jansen /31/) that point processes can be applied for repair. This analysis has been generalized to dependent components but without associated variables. It is important to note that these considerations are strongly related to queuing theory and that they can cover a wide region where fault tree analysis alone is no longer useful.

4.2 Phase type distribution

There is also a second method which can be related to system reliability. This is an algorithmic approach and can be referred to

- computational probability
- matrix geometric methods and to
- phase type distributions (PH-distributions).

It is due to M. Neuts /32/ and his school.

For instance, it could be shown that PH-distributions are very useful for many problems in queuing theory /32/. We only recall that queuing processes are a special type of point processes and Markow renewal processes. They may be used for a number of problems in reliability, e.g. related to repairmen and to computers.

General properties

It has been shown (see Neuts /32/) that the class of PH-distributions is closed under some operations, e.g. under

- finite mixtures of PH-distributions
- convolutions
- formation of maxima and minima
- construction of coherent systems.

Clearly, under these operations with PH-distributions, the resulting distributions are still of phase type, and moreover, it is possible to construct representations of PH-distributions. This is a very interesting development in the region of applied and computational probability. Let us note a few basic concepts.

Definition and some basic properties

We consider a Markow process

$$\{X_t, 0 \leq t < \infty\} \qquad (4.2-1)$$

with a finite number of states labeled 1,...,m+1. We have

$$p_{ij}(s,t) = P\{X_t(w) = j | X_s(w) = i\} \qquad (4.2-2)$$

We may write with (4.2-2) a special case of the Chapman-Kolmogorow-equation characterzing a Markow process. A stochastic process is said to have stationary transition probabilities if for each pair ij the transition probability $p_{ij}(s,t)$ depents only on t-s. This is sufficient for the following discussion. We may write the Chapman-Kolmogorow equation:

$$p_{ik}(t) = \sum_j p_{ij}(s)p_{jk}(t-s) \qquad (4.2-3)$$

with

$$p_{ij}(t) \geq 0, \sum_j p_{ij}(t) = 1 \qquad (4.2-4)$$

With suitable continuity assumptions we have:

$$\lim_{t\to 0} p_{ij}(t) = \begin{cases} 1 & \text{for } i=j \\ 0 & i\neq j \end{cases} \qquad (4.2-5)$$

Assuming that $p_{ij}(t)$ has a derivative $p_{ij}'(t)$ for all $t \geq 0$ and that (4.2-5) holds we may obtain the following relations:

$$q_i = \lim_{t\to 0} \frac{1-p_{ii}(t)}{t} = - p'_{ii}(0) \qquad (4.2\text{-}6)$$

$$q_{ij} = \lim_{t\to 0} \frac{p_{ij}(t)}{t} = p'_{ij}(0) \qquad (4.2\text{-}7)$$

These relations can be used to define the "infinitesimal generators" q_i, q_{ij} of a Markow process.

Let Q be the matrix $[q_{ij}]$, where we use $q_{ii} := -q_i$ as diagonal elements. From (4.2-3) we obtain the backward equation

$$p'_{ik}(t) = \sum_j q_{ij} p_{jk}(t) \qquad (4.2\text{-}8)$$

The q_{ij} determine the $p_{ij}(t)$ uniquely.

This system may be also written in matrix form

$$p'(t) = Q\ P(t) \qquad (4.2\text{-}8)$$

where Q is the infinitesimal generator.

Then we can write a solution

$$P(t) = \exp(Qt) \qquad (4.2\text{-}8)$$

where

$$\exp(Qt) = \sum_{r=0}^{\infty} Q^r \frac{t^n}{r!} \qquad (4.2\text{-}9)$$

Remark: It can be shown that if the eigenvalues of Q are all distinct, we obtain (for(4.2-8)):

$$P(t) = B \begin{pmatrix} e^{\lambda_1 t} & & \\ & \ddots & \\ & & e^{\lambda_r t} \end{pmatrix} C' \qquad (4.2\text{-}10)$$

where BC' = I (4.2-11)

with I identity matrix. (see Cox, Miller /29/).

Infinitesimal generators

We consider a Markow process with the states $\{1,2,\ldots, m+1\}$ and the infinitesimal generator

$$\underline{Q} = \begin{pmatrix} \underline{\Gamma} & \underline{\Gamma}^o \\ \underline{0} & 0 \end{pmatrix} \qquad (4.2\text{-}12)$$

where the mxm - matrix $\underline{\Gamma}$ satisfies

$$\left.\begin{array}{l}\lambda_{ii} < 0 \\ \lambda_{ij} \geq 0\end{array}\right. \quad \text{for } 1 \leq i \leq m$$

we also have this relation:

$$\underline{\Gamma}\ \underline{e} + \underline{\Gamma}^{o} = \underline{0} \qquad (4.2\text{-}13)$$

where e is a unit vector. Moreover, for t = 0 we have

$$(p_1(0),\ldots,p_m(0),p_{m+1}(0)) = (\underline{\alpha},\alpha_{m+1})$$

and

$$\underline{\alpha}\ \underline{e} + \alpha_{m+1} = 1 \qquad (4.2\text{-}14)$$

which is equivalent to (4.2-4).

We assume that all states 1,..,m are transient and that state m+1 is absorbing.

Theorem The probability distribution F of the time until absorption in the state m+1 corresponding to the initial probability vector $(\underline{\alpha}.\alpha_{m+1})$ is given by

$$F(x) = 1 - \underline{\alpha}\ \exp(\underline{\Gamma}x)\cdot\underline{e} \qquad (4.2\text{-}15)$$

where $\underline{\Gamma}$ is the submatrix of the generator $\underline{Q}$ (4.2-12).

Scetch of a proof: We refer to (4.2-8). With initial conditions

$$(p_1(0),\ldots, P_m(0)) = \underline{\alpha}$$

we obtain (due to (4.2-4)) the relation (4.2-15).

Definition: A probability distribution F on $[0,\infty)$ is a distribution of phase type (PH-distribution) iff it is the distribution of the time until absorption in a finite Markow process of the type defined in (4.2-12) (infinitesimal generators). The pair $(\underline{\alpha},\underline{\Gamma})$ is called representation of F.

A few properties of PH-distributions

1. Thes distributions have a jump of height α_{m+1} at x = 0
2. The laplace-Stieltjes transform F*(s) of F(x) is

$$F^*(s) = \alpha_{m+1} + \underline{\alpha}(s\underline{I} - \underline{\Gamma})^{-1}\underline{\Gamma}^{o} \qquad (4.2\text{-}16)$$

3. The noncentral moments μ_i' of F(x) are given by

$$\mu_i' = (-1)^{i}i!(\underline{\alpha}\ \underline{\Gamma}^{-i}\underline{e}) \qquad (4.2\text{-}17)$$

Example: The Erlang distribution of order m (pdf) is:

(pdf) is:

$$f(x) = \frac{\lambda(\lambda x)^{m-1}\ e^{-\lambda x}}{(m-1)!} \qquad (4.2\text{-}18)$$

and has the following representation

$$(\underline{\alpha},\underline{\Gamma}) \qquad (4.2\text{-}19)$$

with $\qquad \underline{\alpha} = (1,0,\dots,0)$

$$\underline{\Gamma} = \begin{pmatrix} -\lambda\lambda & & & & \\ & -\lambda\lambda & & & \\ & & \cdots & & \\ & & & -\lambda\lambda & \\ & & & & -\lambda \end{pmatrix}$$

where $\underline{\Gamma}$ is a mxm-matrix.

Closure properties

It has been indicated that PH-distributions are closed under certain operations. We discuss here:

- convolution and
- construction of coherent systems.

Convolution

Convolution may be used for addition of life lengths. If a failed component is replaced by a spare, the total accumulated life time is obtained by the addition of two life lengths. To express the distribution of the sum of two independent life times (where T_1 has distribution F_1, T_2 distribution F_2) and T_1+T_2 distribution F) we use the convolution

$$F(t) = \int_0^t F_2(t-x)\,d\,F_2(x) \qquad (4.2\text{-}20)$$

Notation: If $\underline{\Gamma}^o$ is an m-vector (4.2-13) and $\underline{ß}$ an n-vector, we denote by $\underline{\Gamma_1}^o\underline{B}^o$, the mxn matrix $\underline{\Gamma_1^o}\ \underline{ß}$, with elements $\underline{\Gamma_1}^o_i\ ß_j$, $1 \le i \le m$, $1 \le j \le n$.

Theorem:

If F(x) and G(x) are both continuons PH-distributions with representations

$$(\underline{\alpha},\underline{\Gamma_1})\ ,\ (\underline{ß},\underline{\Gamma_2})$$

of orders m and n respectively,

then their convolution F*G(x) (see also (4.2-20)) is a PH-distribution with representation $(\underline{\gamma},\underline{L})$ given by

$$\underline{\gamma} = (\underline{\alpha},\ \underline{\alpha}_{m+1}\circ\underline{\beta})$$

$$\underline{L} = \begin{pmatrix} \underline{\Gamma_1} & \underline{\Gamma_1}^o\ \underline{B}^o \\ 0 & \underline{\Gamma_2} \end{pmatrix} \qquad (4.2\text{-}21)$$

Proof: See Neuts /32/. It can be shown, using the Laplace-Stieltjes-transform of F and G (4.2-16) and the product corresponding to a convolution that (4.2-21) holds.

Example: Convolution of Erlang distributions (both of degree 2), but with different failure rates λ_i(i=1,2) and λ_j(j = 1,2). (see (4.2-18)).

F*G(x)

For F(x) we have representation

$$(\underline{\alpha},\underline{\Gamma_1})$$

with $\underline{\Gamma_1} = \begin{pmatrix} -\lambda_1 & \lambda_1 \\ 0 & -\lambda_2 \end{pmatrix}$

and $\underline{Q}_1 = \begin{pmatrix} \underline{\Gamma_1} & \underline{\Gamma_1}^o \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} -\lambda_1 & \lambda_1 & 0 \\ 0 & -\lambda_2 & \lambda_2 \\ 0 & 0 & 0 \end{pmatrix}$.

For G(x) we have representation $(\underline{\beta},\underline{\Gamma_2})$.

Thus F*G is represented by

$$(\underline{\gamma},\underline{L})$$

with $\underline{\gamma} = (\underline{\alpha},\ \alpha_{m+1}\underline{\beta})$

$$\underline{L} = \begin{pmatrix} \underline{\Gamma_1} & \underline{\Gamma_1}\ \underline{B}^o \\ & \\ 0 & \underline{\Gamma_2} \end{pmatrix} = \begin{array}{ccccc} \underline{\Gamma_1} & & \underline{\Gamma_1}^o & \underline{B}^o & \\ -\lambda_1 & \lambda_1 & 0 & 0 & \\ 0 & -\lambda_2 & \lambda_2 & 0 & \\ & & -\lambda_3 & \lambda_3 & \\ & & 0 & -\lambda_4 & \}\underline{\Gamma_2} \end{array} \qquad (4.2\text{-}22)$$

By a representation (see (4.2-20)) and by use of the convolution property of Erlang distributions we can indeed obtain the same matrix $\underline{L}$ (given in (4.2-22)).

Repairable components

The convolution theorem (4.2-21) may be also applied for a repairable component. We have the following structure for the generator of a Markow process:

$$\underline{M} = \begin{pmatrix} \underline{\Gamma}_1 & \underline{\Gamma}_1{}^{o}\underline{B}^{o} \\ \\ \underline{\Gamma}_2{}^{o}\underline{A} & \underline{\Gamma}_2 \end{pmatrix} \qquad (4.2\text{-}23)$$

Without loss of generality, we may assume $\alpha_{m+1} = \beta_{n+1} = 0$. If at time t the Markow process is in the set of states $\{1,2,\ldots,m\}$, the point is covered by an interval with distribution F. A similar consideration holds for sojourns in the set $\{m+1,\ldots,m+n\}$. Transitions between these sets are called renewals. We obtain an alternating renewal process.

Construction of coherent systems

It is sufficient to consider for PH-distributions of life times T_1, T_2 the distribution of min (T_1,T_2) and max (T_1,T_2) (see also (4.1-17)).

Kronecker Product

If $\underline{L}$ and $\underline{M}$ are rectangular matrices of dimensions k_1k_2 and $k_1'k_2'$, their Kronecker product $\underline{L} \otimes \underline{M}$ is defined as the matrix of dimensions $k_1k_1'.k_2k_2'$ written as follows:

$$\underline{L} \otimes \underline{M} = \begin{pmatrix} L_{11}\underline{M} & L_{12}\underline{M} & \ldots. & L_{1k_2}\underline{M} \\ . & & & \\ . & & & \\ . & & & \\ . & & & \\ L_{k_11}\underline{M} & L_{k_12}\underline{M} & \ldots. & L_{k_1k_2}\underline{M} \end{pmatrix} \qquad (4.2\text{-}24)$$

Note that the r.h.s of (4.2-24) is written as a matrix of submatrices (in block partitioned form).

Now, for independent r.v. T_1, T_2 with PH-distributions a theorem will be statet. Let

$$F_{max}(t) = F_1(t)\; F_2(t), \text{ and}$$
$$F_{min}(t) = 1 - \left[1-F_1(t)\right]\left[1-F_2(t)\right] \qquad (4.2\text{-}25)$$

be distributions, corresponding to $\max(T_1,T_2)$ and $\min(T_1,T_2)$ respectively.

Theorem

Let $F_1(t)$ and $F_2(T)$ have representations $(\underline{\alpha},\underline{\Gamma_1})$ and $(\underline{\beta},\underline{\Gamma_2})$ of orders m and n respectively.

(a) Then $F_{max}(t)$ (4.2-25) has the representation $(\underline{\alpha},\underline{L})$ of order mn + m + n, given by

$$\underline{\gamma} = (\underline{\alpha} \otimes \underline{\beta},\; \beta_{n+1}\underline{\alpha},\; \alpha_{m+1}\underline{\beta}) \;,$$

$$\underline{L} = \begin{pmatrix} \underline{\Gamma_1} \otimes \underline{I} + \underline{I} \otimes \underline{\Gamma_2} & \underline{I} \otimes \underline{\Gamma_2}^{0} & \underline{\Gamma_1}^{0} \otimes \underline{I} \\ 0 & \underline{\Gamma_1} & 0 \\ 0 & 0 & \underline{\Gamma_2} \end{pmatrix} \qquad (4.2\text{-}26)$$

where $\underline{I}$ is the unit matrix.

(b) Similarly, $F_{min}(t)$ (4.2-25) has the representation (δ,M) given by

$$\underline{\delta} = (\underline{\alpha} \otimes \underline{\beta})$$
$$\underline{M} = \underline{\Gamma_1} \otimes \underline{I} + \underline{I} \otimes \underline{\Gamma_2} \qquad (4.2\text{-}27)$$

Remarks

We will not go into the details of a proof. But let us note this: For a Markow matrix which is decomposable, a Kronecker product of two Markow matrices represents this decomposition (see Paz /33/). For a proof of this theorem see Neuts /32/. The main step is there to show that $\underline{\Gamma_1} \otimes \underline{I} + \underline{I} \otimes \underline{\Gamma_2}$ cannot be singular. The infinitesimal generators $\underline{\Gamma}$ (see (4.2-12)) are nonsingular matrices.

Examples:

Let us consider the two basic elements of a coherent system. We assume systems with two components where the life times are exponentially distributed, with λ_1,λ_2.

(a) Series system

For min (T_1, T_2) we obtain

$$\begin{aligned} F_{min}(t) &= 1 - (1-F_1(t))(1-F_2(t)) \\ &= 1 - e^{-(\lambda_1+\lambda_2)t} \end{aligned}$$

For a PH-distribution we obtain the following representation (see (4.2-27)):

$$\underline{\Gamma_1} \otimes \underline{I} + \underline{I} \otimes \underline{\Gamma_2} = \lambda_1+\lambda_2$$

(b) Parallel system

For max (T_1, T_2) we obtain

$$F_{max}(t) = F_1(t)\ F_2(t) = (1-e^{-\lambda_1 t})(1-e^{-\lambda_2 t})$$

With (4.2-26) we obtain as representation $(\underline{\gamma}, \underline{L})$, where

$$\underline{L} = \begin{pmatrix} \underline{\Gamma_1} \otimes \underline{I} + \underline{I} \otimes \underline{\Gamma_2} & \underline{I} \otimes \underline{\Gamma_2}^{o} & \underline{\Gamma_1}^{o} \otimes \underline{I} \\ 0 & \underline{\Gamma_1} & 0 \\ 0 & 0 & \underline{\Gamma_2} \end{pmatrix}$$

With (4.2-13) we obtain

$$\underline{\Gamma}_i \cdot \underline{e} + \underline{\Gamma}_i^{o} = 0 \quad (i = 1,2)$$

For exponential distributions, we have

$$-\lambda_i \cdot 1 + \lambda_i = 0 \quad (i = 1,2)$$

The representation is of order $mn + m + n = 1 \cdot 1 + 1 + 1 = 3$.

Finally

$$\underline{L} = \begin{pmatrix} -(\lambda_1+\lambda_2) & \lambda_1 & \lambda_2 \\ 0 & -\lambda_2 & 0 \\ 0 & 0 & -\lambda_1 \end{pmatrix}$$

The same result may be also abtained by a transition matrix and a transition diagram. This transition matrix is closely related to system reliability.

Remarks:

1. It is also possible to generalize this consideration to systems with repairable components.
2. Neuts /32/ mentionend that this result is not yet of computational utility.

4.3 Markow renewal process (MRP)

The Markow renewal process (MRP) is a generalization of Markow processes and of renewal processes. It is one of the best known processes with non-Markowian behavior.

It is possible to evaluate for a system suitable measures of effectiveness (reliability, availability, maintainability) using Markow renewal processes (MRP). This can be done with techniques known partly from Markow processes. We discuss a few basic concepts, show relations to fault tree analysis and stopping times and mention a few techniques for evaluation. But also problems which are not suitable for fault tree analysis can be dealt woth MRP.

Notations and assumptions:

Suppose we have defined for each $n \in \mathbb{N}$, a random variable taking values in a finite set E and a random variable T_n taking values in $\mathbb{R}_+ = [0,\infty)$ such that

$$0 = T_o \leq T_1 \leq T_2 \leq \ldots$$

The set E (for our purpose a finite set) gives the possible states, T_n gives the sojourn times $(n = 0,1,2,\ldots)$.

Def. The stochastic process $X_t = \{X_n, T_n;\ n \in \mathbb{N}\}$ is called a Markow renewal process with state space E, provided that

$$P\{X_{n+1} = j,\ T_{n+1}-T_n \leq t | X_o,\ldots,X_n;\ T_o,\ldots,T_n\} = P\{X_{n+1} = j,\ T_{n+1}-T_n \leq t | X_n\} \qquad (4.3\text{-}1)$$

for all $n \in \mathbb{N}$, $j \in E$ and $t \in \mathbb{R}_+$.

Remark: Markow renewal processes (MRP) are also closely related to semi-Markow-processes.

A few Properties

We shall require that X_t is time homogeneous, i.e. for any $i,j \in E$, $t \in \mathbb{R}_+$

$$P\{X_{n+1} = j,\ T_{n+1} - T_n \leq t | X_n = i\} = Q(i,j,t) \qquad (4.3\text{-}2)$$

to be independent of n.

The family of probabilities (defined in (4.3-2))

$$Q = \{Q(i,j,t)\ ;\ i,j \in E,\ t \in \mathbb{R}_+\} \qquad (4.3\text{-}3)$$

is called a semi-Markow kernel (over E).

Properties: For each pair (i,j) the function $t \to Q(i,j,t)$ has all properties of a distribution. But we note that

$$P(i,j) = \lim_{t\to\infty} Q(i,j,t) \qquad (4.3\text{-}4)$$

is generally not equal to 1. Here the relations (4.3-5) hold:

$$P(i,j) \geq 0, \quad \sum_{j \in E} P(i,j) = 1 \qquad (4.3\text{-}5)$$

This means that P(i,j) are transition probabilities of a Markow chain.

Characterization of a MRP

A MRP can be completely characterized by

(a) the initial distribution

$$P(X_o = j) = \Pi_j \qquad j \in E \qquad (4.3\text{-}6)$$

(b) the semi-Markow kernel $Q(t) = (Q(i,j,t))$ (see (4.3-2)), with requirements (4.3-4) and (4.3-5).

Other characterizations are useful for evaluation:

Distribution of sojourn times: For this distribution we define

$$G(i,j,t) := \frac{Q(i,j,t)}{p(i,j)} \qquad (4.3\text{-}7)$$

if $p(i,j) \neq 0$, otherwise we set $Q(i,j,t)/p(i,j) = 1$.

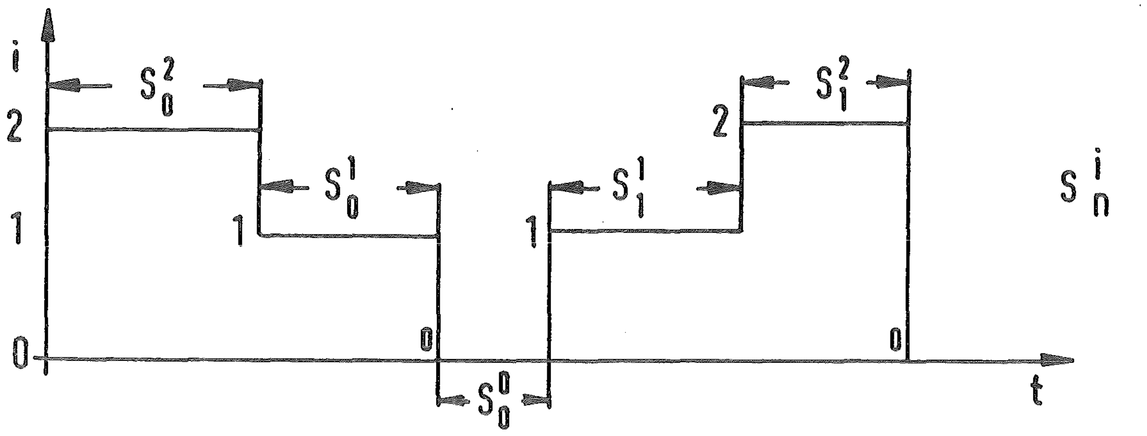
Thus, a MRP can be characterized as follows:

(a') the initial distribution

(b') a matrix of distributions G(i,j,t)

(c') a transition matrix P(i,j)

Example



A realization for a MRP with three states (j=0,1,2).

(a') The initial distribution is

$\Pi_2 = 1, \quad \Pi_1 = 0, \quad \Pi_o = 0$

(b') In the present state i, a random mechanism "chooses" the next state j according to the transition matrix P(i,j), see (4.3-4).

(c') For the present state i, a different random mechanism "determines" the sojourn time S_n^i in this state according to the matrix of distributions G(i,j,t), see (4.2-7).

Special cases

It can be seen that Markow processes and renewal processes are special cases of MRP.

1. If all T_n are equal to 1, we need only a transition matrix P(i,j). Thus we have a Markow chain.
2. If all T_n are exponentially distributed, we have a Markow process.
3. If the state space E consists of a single point, we have a renewal process.

Markow renewal function

In relation to the renewal function (see (2.2-4)) it is possible to introduce a Markow renewal function.

Let j be fixed, and define S_o^j, S_1^j, ... as the successive T_n for which X_n=j.

(See also Fig.). Then $S^j = \{S_n^j \; ; n \in \mathbb{N}\}$ is a (possibly delayed) renewal process. The number of renewals during $[0,t]$ of this renewal process is $N_{ij}(t)$. Now we obtain the conditional expected number of type j events in $[0,t]$ under the condition that this renewal process started with an event of type i at t = 0.

$$H(i,j,t) = E(N_{ij}(t)|X_o=i) \tag{4.3-8}$$

This can be also related to a renewal density:

$$h(i,j,t)\Delta t = P\{\text{event of type } j \text{ in } (t,t+\Delta t)|\text{event of type } i \text{ at } t=0\} \tag{4.3-9}$$

It is important to note that the functions H(i,j,t) are Markow renewal functions, and the collection $H = \{H(i,j,\cdot);\ i,j\varepsilon E\}$ is called a Markow renewal kernel. By an integral equation (Markow renewal equation) this can be related to the semi-Markow kernel.

Result: It can be shown that a fault tree with components which are represented by alternating renewal processes can be represented as a whole using a Markow renewal process.

It is now evident that the Markow renewal functions H(i,j,t) (and their respective renewal densities h(i,j,t)) can be interpreted as expected number of failures/repairs (and their respective failure intensities/ repair intensities). See also (2.3-10) to (2.3-13).

Analysis:

Of course, for practical problems the evaluation of H (and Q) in the Laplace domain (similar to (2.2-7)) is preferable. For evaluation of a MRP various methods can be used. We name only a few:

(a) The method of stages: The device of stages is a method of representing a non-exponentially distributed state by a combination of stages each of which is exponentially distributed. Any distribution with a rational Laplace transform can be represented exactly. Other distributions can be approximated. This has some relation to the method of phase type distributions (Cox /11/, Neuts /32/).

(b) Supplementary variables: A sufficient number of supplementary variables is added to obtain a Markow process. This is direct, but may be cumbersome for evaluation (Cox and Miller /29/).

(c) Inbedded Markow process: We consider a suitable discrete set of time points so that the new process is Markow at a series of time points. This involves some requirements. But it is especially useful for steady state results. (Cox and Miller /29/, König and Stoyan /28/).

There are also other methods available which come frequently from queuing theory (König and Stoyan /28/, Gnedenko and Kowalenko /34/).

Summary of section 4

As a summary of all the stochastic processes mentioned we are giving a table listing

- the type of process (also referring to sections of this paper)
- the type of component or system which can be modelled
- the distributions which may be used with this process (life time and repair time distributions)
- a few topics belonging th the required background including a reference
- an estimated degree of difficulty.

Similar tables may be found in Corazza /35/, and in König, Stoyan /28/.

stochastic process (section)	can be used for modelling	distributions	background	degree of difficulty
renewal process (2.2)	spare parts reservation with negligible repair time	arbitrary (life time distr.)	renewal theory Cox /11/	medium
alternating renewal process	repairable components (without restriction of repair time)	arbitrary life and repair time	renewal theory Cox /11/	medium
Markow process	systems with arbitrary structure (practical limits for medium/large size systems)	only exponential	Laplace transform, eigenvalue problems Corazza /35/ Cox,Miller /29/ Kemeny,Snell /36/	medium
Markow renewal process (4.3)	systems with arbitrary structure (practical limits for medium size systems)	exponential failure distr., arbitrary repair distr.	Laplace-Stiltjes Transform, Inversion, e.g. imbedded Markow chain Corazza /35/ Cinlar /37/	high
Point process (4.1)	coherent structures (mostly for methodological considerations)	arbitrary	theory of point processes, stopping times Brémaud /24/	very high
phase type distributions (4.2)	coherent structures (for methodological consideration), very good for queuing processes	PH-distribution e.g. Erlang-distr.	matrix analytic methods, Markow process Neuts /32/	high

5. Systems with Phased Mission

5.1 Introduction

Until now we discussed systems which have the same configuaration during the whole life time. If we have, however, a system with a phased mission, its configurations may change during consecutive periods (called phases). Reliability and performance analysis requires the use of a (generalized) multistate structure function and the concept of association (see Barlow /2/).
It is possible to give bounds for unavailability. It is interesting to note that there is also a criterion showing the admissibility of phased structure functions for these systems. This can be based on some algebraic properties of the so called functional dependence (see Meyer /38/).

It will be sufficient to consider here systems having two states for each component. Fore more general information see Esary and Ziehms /39/, A. Pedar and V. Sarma /40/).

5.2 Discussion of a phased mission

We consider the system of Fig. given as block diagram. It has different structures in the three phases of its mission (see /39/).

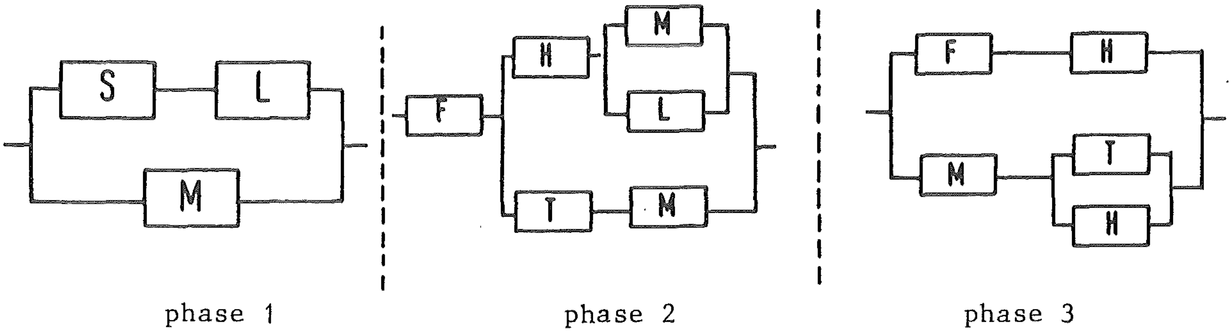


Fig. 5.1 System with phased mission

For this system we obtain as minimal cuts:

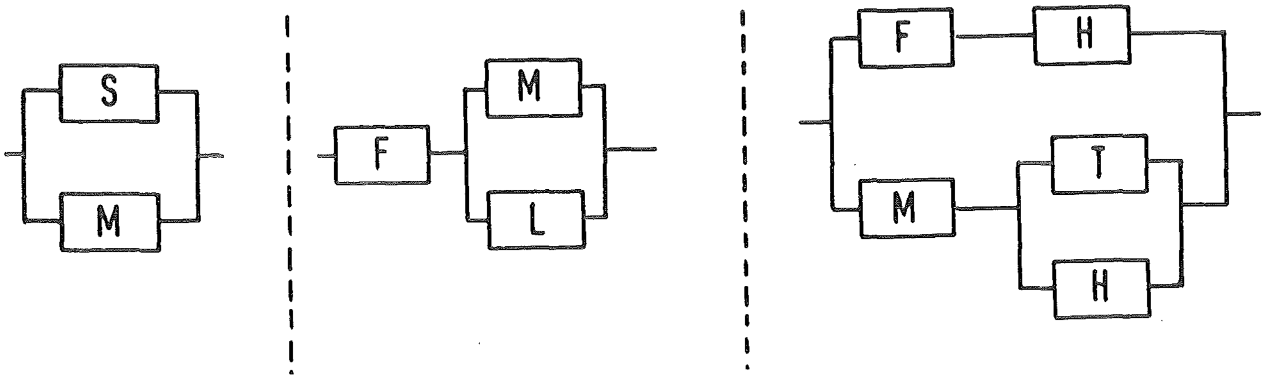
Phase	Minimal cuts
1	{M,L} , {M,S}
2	{F} , {H,M} , {H,T} , {M,L}
3	{F,M}, {H,M}, {H,T}

Simplification of a system

A minimal cut in a phase can be deleted (without loss of information) if it contains a minimal cut of a later phase. This is similar to absorption. But it would not refer to deleting of a minimal cut regardless of time ordering. Thus we obtain the following reduced list ("after cancellation") of cut sets:

Phase	Cuts	cancelled cuts
1	{M,S}	{M,L}
2	{F} , {M,L}	{H,M} , {H,T}
3	{F,M}, {H,M}, {H,T}	no cancellation possible

This can be also given as a simplified block diagramm:



phase 1 phase 2 phase 3

Fig. 5.2 System after cancellation

An equivalent representation is by a structure function Φ_i referring to phase i.

Note: We write all variables x_{Mi} and structure functions for a success tree. Later on, we also introduce a corresponding fault tree.

x_{Mi}(i=1,2,3) refers to the success of component M in phase i. If for a phase j < i, M would be failed, it could not be successful in phase i.

We obtain:

$$\Phi_1(\underline{x}_1) = x_{M1} + x_{S1}$$

$$\Phi_2(\underline{x}_2) = x_{F2}(x_{M2} + x_{L2}) \qquad (5.2\text{-}1)$$

$$\Phi_3(\underline{x}_3) = x_{F3}\, x_{H3} + x_{M3}(x_{T3}+x_{H3})$$

We obtain as probability that this system is operative for the whole mission

$$P_{\text{system}} = P\{ \prod_{j=1}^{n} \Phi_j(\underline{x}_j) = 1\} \qquad (5.2\text{-}2)$$

or

$$E\{ \prod_{j=1}^{n} \Phi_j(\underline{x}_j) = 1\} \leq \prod_{j=1}^{n} E\{\Phi_j(\underline{x}_j) = 1\} \qquad (5.2\text{-}3)$$

(see Esary /39/).

This is an example for a "structure based" capability function, i.e. a function which can be related to structure functions Φ_i (see Meyer /38/).

Kernel of a Boolean mapping

Now we introduce some further considerations which can be used for a methodology of systems with phased missions.

Let Φ_j be a Boolean mapping, from B to A:

$$\Phi_j : B \rightarrow A \qquad (5.2\text{-}4)$$

Then the kernel of Φ_j is the set M_j of elements in B which Φ_j maps onto 1 in A. This can be written:

$$M_j = \{p|\ \Phi_j(\underline{p}) = 1\} \qquad (5.2\text{-}5)$$

Example: The kernel M_1 of Φ_1 is

$$M_1 = \{(x_{M1},x_{S2})|\Phi_1(x_{M1},x_{S1}) = 1\}$$

$$= \{x_{M1},x_{S1}\}$$

Note: p refers to the variables of Φ_j.

Application to our system:

We obtain as kernels:

$$M_1 = \{x_{M1}, x_{B1}\} \qquad (5.2\text{-}6)$$

$$M_2 = \{x_{F2} x_{M2}, x_{F2}\, x_{L2}\}$$

$$M_3 = \{x_{F3}\, x_{H3}, x_{M3}\, x_{T3}, x_{M3}\, x_{H3}\}$$

By a Cartesian product of these kernels

$$M_1 \times M_2 \times M_3$$

we obtain all success trajectories of our system. This can be rewritten:

$$M_1 \times M_2 \times M_3 \qquad (5.2\text{-}7)$$

$$= \{x_{M1}, x_{S1}\} \times \{x_{F2}\, x_{M2}, x_{F2}\, x_{L2}\}$$

$$\times \{x_{F3} x_{H3}, x_{M3}\, x_{T3}, x_{M3}\, x_{H3}\}$$

This Cartesian product can be also given as a tree. In this tree each path from left to right is a single term of the Cartesian product. Each term is a success trajectory.

For example:

M1 • F2 M2 • F3 H3

is a success trajectory. But failure of M1 and S1 would lead to system failure.

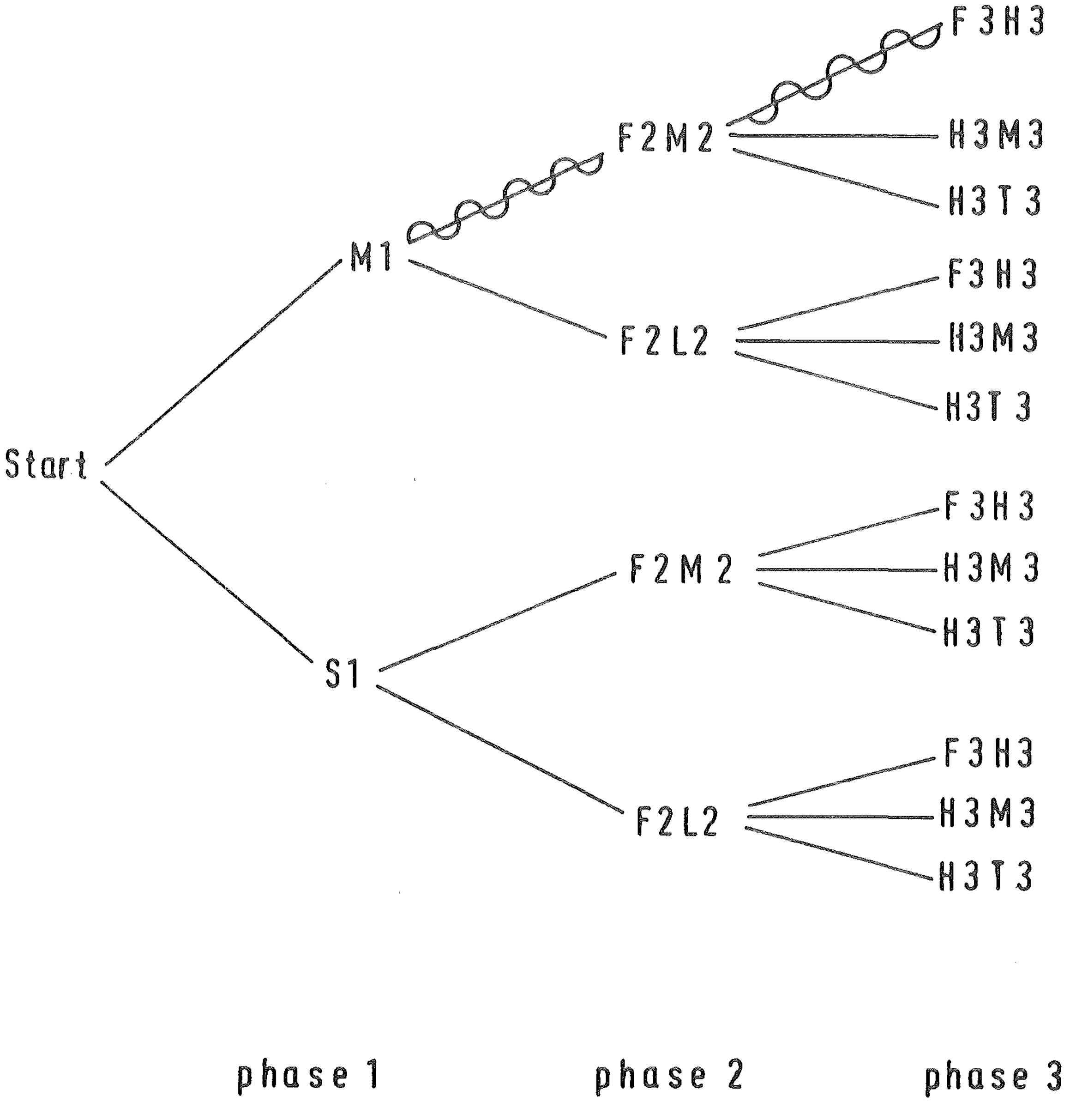


Fig. 5.3 Tree for a system with phased missions

(〰 success trajectory M1 ∘ F2 M2 ∘ F3 H3)

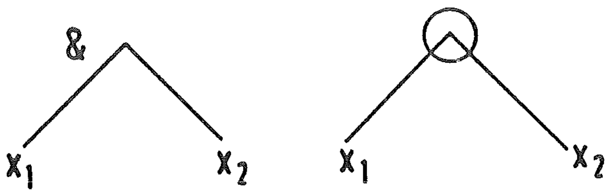
Cartesian product:

$$
\begin{aligned}
& M_1 \text{ x } M_2 \text{ x } M_3 \\
& = \{ x_{M1}\, x_{F2}\, x_{M2}\, x_{F3}\, x_{H3},\; x_{M1}\, x_{F2}\, x_{M2}\, x_{H3}\, x_{M3},\; x_{M1}\, x_{F2}\, x_{M2}\, x_{H3}\, x_{T3}, \ldots \\
& \quad \ldots,\; x_{S1}\, x_{F2}\, x_{L2}\, x_{H3}\, x_{T3} \} \qquad (5.2\text{-}8)
\end{aligned}
$$

Success tree and fault tree

We may also use a success tree or a fault tree for representation of (5.2-1), (5.2-2) or (5.2-7).

Here the symbols



denote conjunction / disjunction of x_1, x_2.

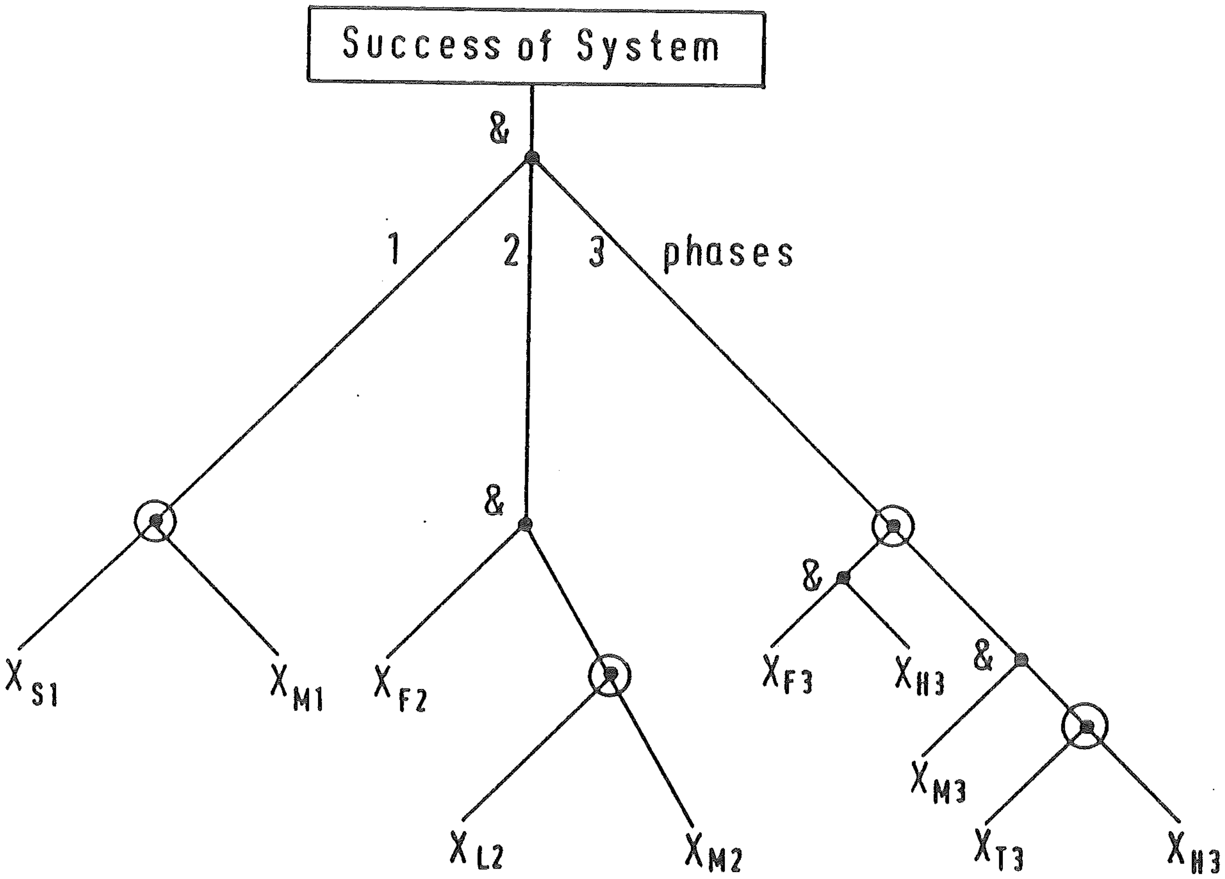


Fig. 5.4 Success tree with three phases

Note:

x_{ke} := component k is intact in phase 1.

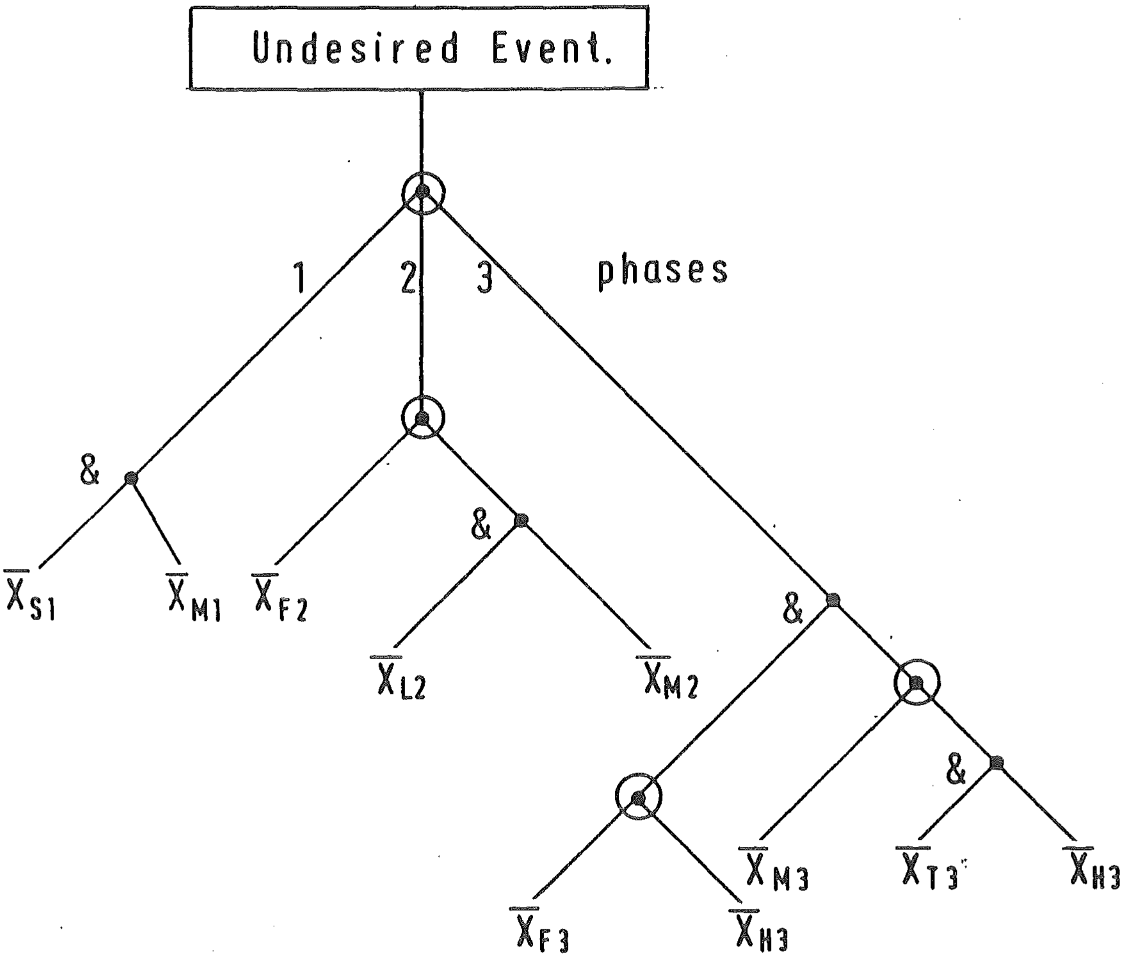


Fig. 5.5 Corresponding fault tree with three phases

Note:

$\bar{x}_{kl}$:= component k failed in phase 1.

5.3 A System which is not Structure Based

By a simple restriction we may obtain a system which cannot be evaluated by a fault tree. We call such a system "not structure based".

For the system of section 5.2. (Fig. 5.3, success paths) we make the following restriction.

Restriction: If in phase 1, the success path went over M1, then in phase 2, F2 M2 is no longer a part of the success path. But F2 L2 still remains.

If in phase 1, the success path went over S1, then in phase 2, F2 M2 is a part of the success path.

Let us show a diagram for this situation.

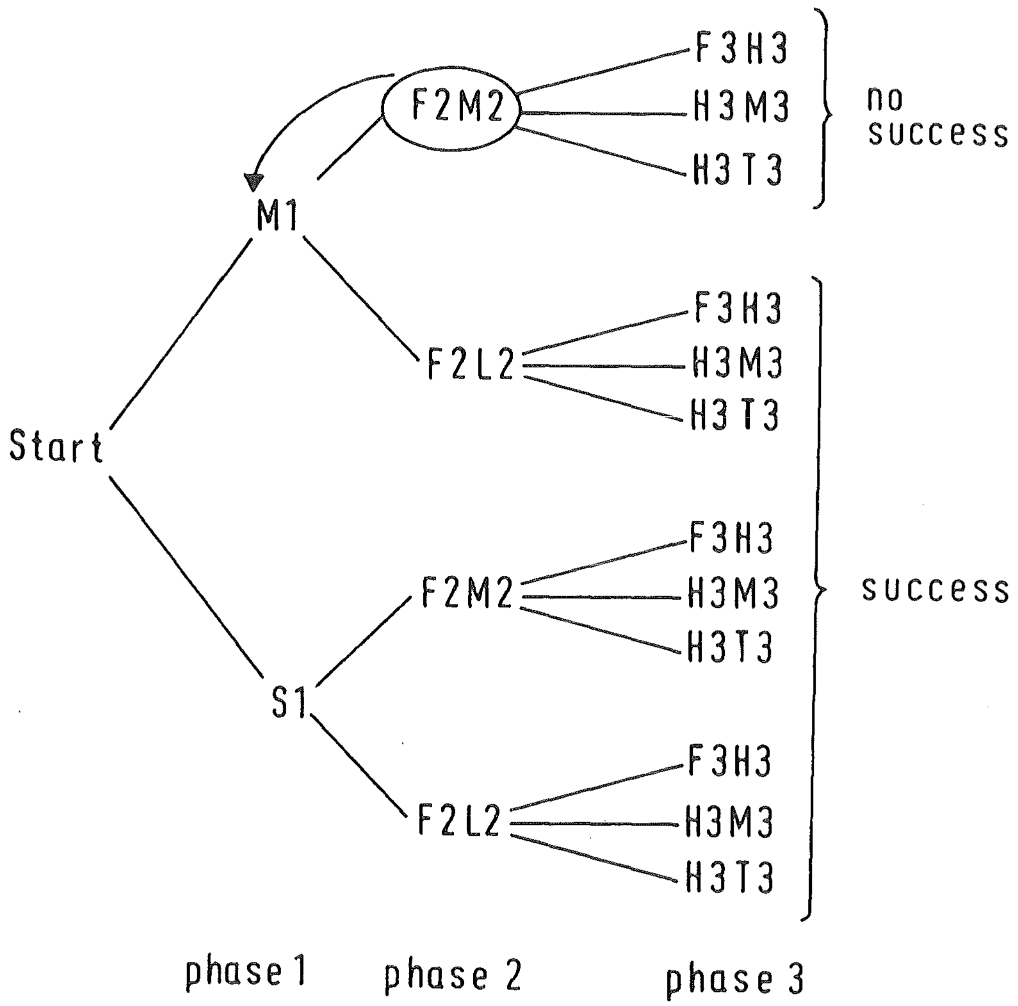


Fig. 5.6 Tree for a system with phased missions

(If a path goes over M1 and F2 M2 no success is obtained, but if a path goes over M1 and F2 L2 success is obtained).

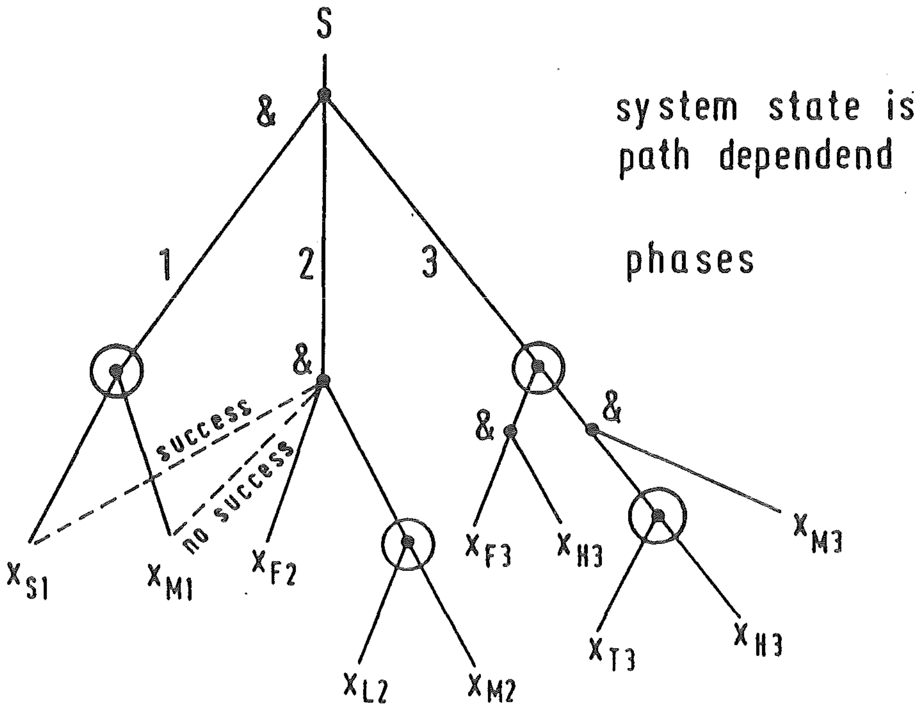


Fig. 5.7 A system which is not structure based

Discussion:

We note

$x_{M1}\ x_{F2}\ x_{M2}\ x_{F3}\ x_{H3}$ no success

but $x_{S1}\ x_{F2}\ x_{M2}\ x_{F3}\ x_{H3}$ is a success.

Moreover, M1 may not be deleted, since then the success path

$x_{M1} \cdot x_{F2}\ x_{L2} \cdot x_{F3}\ x_{H3}$

would vanish. Here the Boolean structure is no longer valid to represent the situation. An equivalent statement holds for a system in terms of failure. The top event cannot be defined by vertices which depend only on predecessors. Thus clearly our fault tree definition is violated. But methods developed for systems which are not structure based can be applied (see Meyer /38/).

5.4 A second System which is not Structure Based

We assume a system which has a given task, e.g. as processor of a computer system (see also Schriefer et al. /41/, on a reliable microcomputer-based LMFBR protection system). An important requirement to this system is that its average throughput τ_{av} over a aperiod T (utilization period) has to be above a prespecified level.

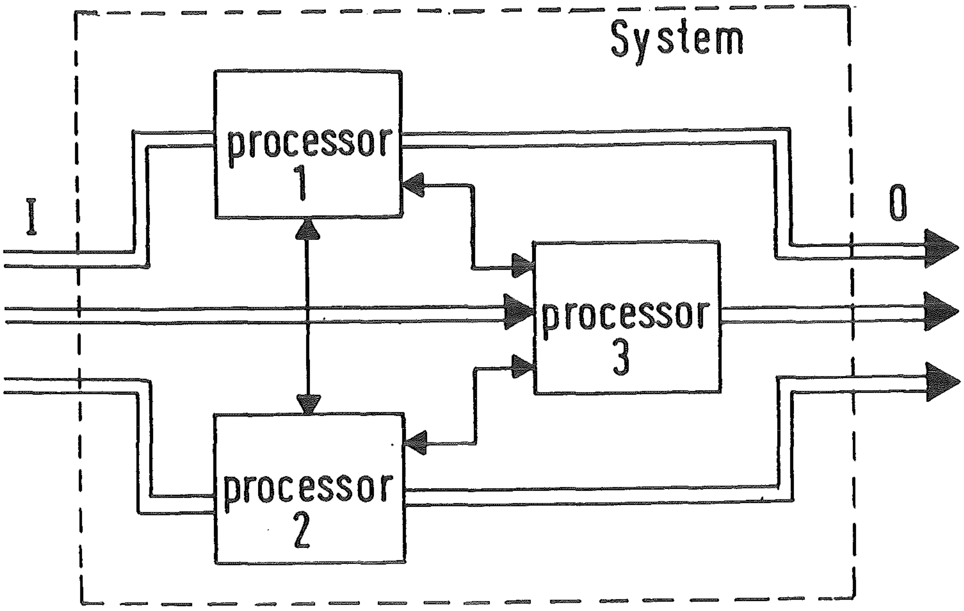


Fig. 5.8 Redundant Processor (Triplicated configuration)

The states of the system (Fig. 5.8) are the following:

State	description	throughput
2	all processors fault free	τ
1	1 processor faulty	$\tau/_2$
0	2 or more processors faulty	0

The utilization period consists of n phases. It is e.g. required that the average throughput is

$$\tau_{av} \geq \frac{\tau}{2} \quad .$$

Example

Let n=3, T consist of 3 phases. One possible state trajectory is (2,1,2), where the general form is

$$u = (q_1, q_2, q_3) \tag{5.4-1}$$

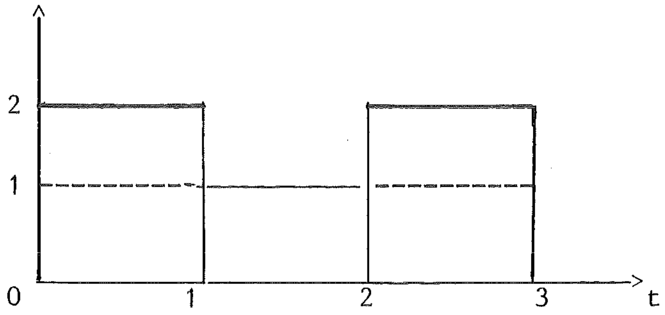


Fig. 5.9 Trajectory u = (2,1,2)

We obtain (from 3 phases):

$$\tau_{av} = \frac{\tau_1 + \frac{\tau_2}{2} + \tau_3}{3} = \frac{5}{6}\,\tau > \frac{\tau}{3} \tag{5.4-2}$$

It is possible to define the following accomplishment levels (a_i) which are related to values of τ_{av}:

accomplishment level a_i	average throughput
a_2	$\tau_{av} \geq \frac{5}{6}\,\tau$
a_1	$\frac{5}{6}\,\tau > \tau_{av} \geq \frac{\tau}{2}$
a_0	$\frac{\tau}{2} > \tau_{av}$

Note that for the levels a_2, a_1 we have system success, while for a_o we have no system success (not sufficient throughput). To each combination of states (trajectory) an accomplishment level a_i can be related, using a capability function $\gamma_s(u)$.

This is (partly) shown in Table 5.4-1.

States in phases 1,2,3	Values of $\gamma_s(u)$	τ_{av}	
(2,2,2)	a_2	$\tau_{av} \geq \frac{5}{6}\tau$	
(2,2,1)	a_2		
(2,2,0)	a_1	$\frac{5}{6}\tau > \tau_{av} \geq \frac{\tau}{2}$	
(2,1,2)	a_2	$\tau_{av} \geq \frac{5}{6}\tau$	1)
(2,1,1)	a_1	$\frac{5}{6}\tau > \tau_{av} \geq \frac{\tau}{2}$	
(2,1,0)	a_1		
⋮	⋮	⋮	
(0,0,2)	a_0	$\frac{\tau}{2} > \tau_{av}$	
(0,0,1)	a_0		
(0,0,0)	a_0		

1) see equ. (5.4-2)

Table 5.4-1 Accomplishment levels

Functional dependence

We can - in analogy to the kernel of a Boolean mapping, (5.2-5) - define the set of states which correspond to a given accomplishment level a_i. Thus we obtain (inverting the capability function $\gamma_s(u)$) the following set:

$$\begin{aligned} R &= \gamma_s^{-1}(a_2) \\ &= \{(2,2,2),\ (2,2,1),\ (2,1,2),\ (1,2,2)\} \end{aligned} \qquad (5.4\text{-}3)$$

The set R of elements $u = (q_1 q_2, q_3)$

(or trajectories) is mapped by the capability function $\gamma_s(u)$ on a_2.

Here we obtain an important conclusion.

<u>If</u> a state trajectory $u = (q_1 . q_2, q_3)$ belongs to the subset with a_2 <u>and</u> <u>if</u> we know that $q_2 = 1$, <u>then</u> we can infer that $q_1 \neq 1$. (5.4-4)

This follows from (5.4-2), (5.4-3) and Fig. 5.9, or from Table 5.4-1.

This means: Knowledge of a state of this system at the end of a phase increases our knowlegde of the previous phase. Similar conderations will be made for a refined capability function γ_s.

Refinement of capability function γ_s

It is frequently useful to have for different values of τ_{av} different accomplishment levels. This can be done as follows:

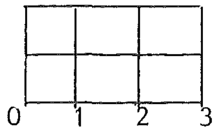
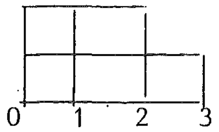
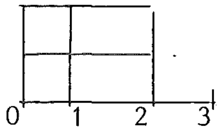
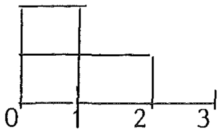
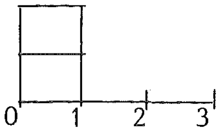
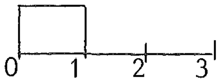
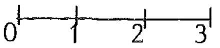
(q_1,q_2,q_3)	trajectory	τ_{av}	a_i
(2,2,2)		τ	6
(2,2,1)		$\frac{5}{6}\tau$	5
(2,2,0)		$\frac{2}{3}\tau$	4
(2,1,0)		$\frac{\tau}{2}$	3
(2,0,0)		$\frac{\tau}{3}$	2
(1,0,0)		$\frac{\tau}{6}$	1
(0,0,0)		0	0

Table 5.4.-2 A refined capability function γ_s

Possible trajectories

We now present all possible trajectories in Table 5.4-3 with the corresponding τ_{av} and a_i $(i=0,1,...,6)$. A representation using a tree would also be posible. However, due to τ_{av} we will have operations which have (in contrast to fault trees!) an inherent memory.

Functional dependence

Recall that we can, in analogy to the kernel of a Boolean mapping (5.2-5) define the set of all states which correspond to a given accomplishment level a_i (see Table 5.4-3). If a_i is an accomplishment level, then the probability that the system S performs at level a_i is given

$$p_s(a) = P(\{u|\gamma_s(u) = a\}) = P(\gamma_s^{-1}(a)) \qquad (5.4\text{-}5)$$

The inverse image $\gamma_s^{-1}(a)$ is referred to as trajectory set of a.
It evidently relates to the kernel (5.2-5).
Example: T a=5 belongs the trajectory set: {(2,2,1), (2,1,2), (1,2,2)}.

t_o q_1 =	2									1									0								
t_1 q_2 =	2			1			0			2			1			0			2			1			0		
t_2 q_3 =	2	1	0	2	1	0	2	1	0	2	1	0	2	1	0	2	1	0	2	1	0	2	1	0	2	1	0
t_3 τ_{av}	1	$\frac{5}{6}$	$\frac{2}{3}$	$\frac{5}{6}$	$\frac{2}{3}$	$\frac{1}{2}$	$\frac{2}{3}$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{5}{6}$	$\frac{2}{3}$	$\frac{1}{2}$	$\frac{2}{3}$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{2}{3}$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{6}$	0
a_i	6	5	4	5	4	3	4	3	2	5	4	3	4	3	2	3	2	1	4	3	2	3	2	1	2	1	0

Table 5.4-3 Possible Trajectories

Now we give a number of examples which illustrate functional dependence (compare also (5.4-4)).

States observed	accomplishment level	additional knowledge
$(q_3 = 1) \wedge$	$(a = 5) \rightarrow$	$(q_1 = 2) \wedge (q_2 = 2)$
$(q_3 = 0) \wedge$	$(a = 4) \rightarrow$	$(q_1 = 2) \wedge (q_2 = 2)$
$(q_2 = 2) \wedge (q_3 = 1) \wedge$	$(a = 4) \rightarrow$	$(q_1 = 1)$
$(q_2 = 2) \wedge (q_3 = 0) \wedge$	$(a = 3) \rightarrow$	$(q_1 = 1)$
$(q_2 = 1) \wedge (q_3 = 1) \wedge$	$(a = 3) \rightarrow$	$(q_1 = 1)$
$(q_2 = 1) \wedge (q_3 = 0) \wedge$	$(a = 2) \rightarrow$	$(q_1 = 1)$
$(q_3 = 2) \wedge$	$(a = 2) \rightarrow$	$(q_1 = 0) \wedge (q_2 = 0)$
$(q_3 = 1) \wedge$	$(a = 1) \rightarrow$	$(q_1 = 0) \wedge (q_2 = 0)$

Table 5.4-4 Examples for dependence

Note: Different representations for functional dependence are also:

$$\begin{aligned} &(q_3 = 1) \wedge (a = 4) \rightarrow (q_1 \neq 0) \wedge (q_2 \neq 0) \\ &(q_3 = 0) \wedge (a = 3) \rightarrow (q_1 \neq 0) \wedge (q_2 \neq 0) \end{aligned} \qquad (5.4\text{-}6)$$

All relations can be derived from Table 5.4-3.

Probability of system performance at level a_i

We are not only interested in availability or failure probability but in the probability that a system performs at level a, i.e. $p_s(a)$. Only if $\gamma_s(u)$ is "structure based" (i.e. belongs to a system without functional dependence), we obtain for $p_s(a)$ the expectation of the structure function (see (3.1-3)). Assume that the user is interested in a performance level, e.g. corresponding to the average throughput τ_{av}, where the average is taken over a utilization period T. We identify system success with a specified minimum τ_{av}, eg. $\tau_{av} \geq \frac{\tau}{2}$. Then capability function is given as

$$\gamma_s(u) = \begin{cases} 1 \text{ if } & \frac{1}{h} \int_o^h \tau(u(t))dt \geq \frac{\tau}{2} \\ 0 & \text{otherwise} \end{cases} \tag{5.4-7}$$

γ_s will generally not admit a formulation which bases on a structure function. We can either find γ_s by integration (5.4-7) or by summation (Table 5.4-3). These operations have an inherent memory (see also J.F. Meyer /38/).

We come to the following conclusions:

Result 1

It could be shown that the inadmissibility of a Cartesian Product Representation is equivalent to functional dependence. We recall that functional dependence can be defined as an increased knowledge on states which could not directly be observed.

Result 2

For systems with Functional Dependence methods of reliability analysis and performance analysis are required which clearly go beyond fault tree analysis.

Conclusion

We give in Table 5.4-5 some limits of fault tree analysis. For details the corresponding sections should be consulted.

Fault Tree Analysis	Comments	Type of Limit	Other Methods Needed
System represented by fault tree (section 1,2,3)	equivalent: combinational circuit (Weber /42/)	LOGIC	Sequential circuit: contradicts fault tree definition (p.1), probabilistic automata theory (Paz /33/).
on component level: alternating renewal process on system level: point process, MRP (section 4)	components - without repair - with repair - with inspection - with statistical dependence (see limitations for $\bar{R}_s$, p. 20-21)	PROBA-BILISTIC	A system with time sequence of events where average amount of radioactive release has to be evaluated: no representation by fault tree possible, MRP with nonlinear cost functions.
system with phased mission (section 5.1, 5.2)	fault tree representation possible (absence of functional dependence)	ALGE-BRAIC	Systems with functional dependence: Result 1 and 2 of section 5.4, e.g. FTCS with capability function, use of stochastic processes or simulation (Meyer /38/).

Table 5.4-5 Some Limits of Fault Tree Analysis

References

/1/ MURCHLAND, J.D., WEBER, G.
A Moment Method for the Calculation of a Confidence Interval for the Failure Probability of a System, Proceedings 1972 Annual Symposium on Reliability, San Francisco, pp. 565, 1972

/2/ BARLOW, R.E., PROSCHAN, F.
Statistical Theory of Reliability and Life Testing, Probability Models, Holt, Rinehart and Winston, New York, 1975

/3/ CHU, T.L., APOSTOLAKIS, G.
Methods for Probabilistic Analysis of Noncoherent Fault Trees, IEEE Trans. Reliability, Vol. R.29, pp. 354-360, 1980

/4/ KOHAVI, Z.
Switching and Finite Automata Theory (2nd Ed.), Mc Graw Hill Book Company, New York, 1978

/5/ BARLOW, R.E., FUSSELL, J., SINGPURWALLA, N. (Ed.)
Reliability and Fault Tree Analysis, SIAM, Philadelphia, PA, 1978

/6/ FUSSELL, J.B., VESELY, W.E.
A New Methodology for Obtaining Cut Sets for Fault Trees, Trans. American Nuclear Society, Vol. 15. pp. 262-263, 1972

/7/ BENNETTS, R.G.
On the Analysis of Fault Trees IEEE Trans. Reliability, Vol. R-24 pp. 175-185, 1975

/8/ NAKASHIMA, K.
Studies on Reliability Analysis and Design of Complex Systems, PhD Thesis, Kyoto University, Kyoto, 1980

/9/ NAKASHIMA, K., HATTORI, Y.
An Efficient Bottom-Up Algorithm for Enumeration Minimal Cut Sets of Fault Trees, IEEE Trans. Reliability, Vol. R-28, pp. 353-357, 1979

/10/ NELSON, R.J.
Simplest Normal Truth Functions, J. Symbolic Logic, Vol. 20, pp. 105-108, 1954

/11/ COX, D.R.
Renewal Theory, Methuen & Co. Ltd., London, 1962

/12/ HÖFELE-ISPHORDING, U.
Zuverlässigkeitsrechnung, Einführung in ihre Methoden, Springer Verlag, Berlin, 1978

/13/ DOOB, J.L.
Stochastic Processes, Wiley and Sons, New York, 1953

/14/ MURCHLAND, J.D.
Fundamental Concepts and Relations for Multi-State Reliability in /5/, pp. 581-618

/15/ JOKELA, S.
The Availability and Reliability of Complex Systems, Electrical Engineering Laboratory, Report 15, Espoo, Finland, 1976

/16/ NAGEL, K., WEBER, G.
Importanzkenngrößen für die Zuverlässigkeitsanalyse von Systemen, in VDI-Bericht 395, Tagung "Technische Zuverlässigkeit", Nürnberg, S. 145, 1981

/17/ BOROWKOW, A.A.
Wahrscheinlichkeitstheorie, Birkhäuser-Verlag, Basel, 1976

/18/ OLMOS, J., WOLF, L.
A Modular Representation and Analysis of Fault Trees, Nuclear Engineering and Design, Vol. 48, pp. 531-561, 1978

/19/ MURCHLAND; J.D.
Comment on "A Time Dependent Methodology for Fault Tree Evaluation", Nuclear Engineering and Design, Vol. 22, pp. 167-169, 1972

/20/ CLAROTTI, C.A.
Limitations of Minimal Cut Set Approach in Evaluating Reliability of Systems with Repairable Components, IEEE Trans. Reliability, Vol. R-30, pp. 335-338, 1981

/21/ VESELY, W. E.
A Time Dependent Methodology for Fault Tree Evaluation, Nuclear Engineering and Design. Vol. 13, pp. 337.360, 1970

/22/ CALDAROLA, L., WICKENHÄUSER, A.
The Boolean Algebra with Restricted Variables as a Tool for Fault Tree Modularization, KfK 3190/EUR 7056e, Karlsruhe 1981

/23/ ROSENTHAHL, A.
Decomposition Methods for Fault Tree Analysis, IEEE Trans. Reliability, Vol. R-29, pp. 136-138, 1980

/24/ BRÉMAUD, P.,
Point Processes and Queues: Martingale Dynamics, Springer Verlag, New York, 1981

/25/ BIRNBAUM, Z.W., ESARY, J.D., MARSHALL, A.W.
Stochastic Characterization of Wearout for Components and Systems, Ann. Math. Statist. 37, pp. 816-825, 1966

/26/ GREENWOOD, P.
Point Processes and System Lifetimes in "Stochastic Differential Systems" (Proc. of 3rd IFIP-WG7/1 Working Conference Sept. 1980), Lecture Notes in Contr. and Inform. Sciences No 36, Springer Verlag, Berlin, S. 56-60, 1981

/27/ ARJAS, E.
The Failure and Hazard Processes in Multivariable Reliability Systems, Math. of Operations Research 6, pp. 551-562, 1981

/28/ KÖNIG, D., STOYAN, D.
Methoden der Bedienungstheorie, Akademie-Verlag, Berlin, 1976 (Vieweg-Braunschweig 1976)

/29/ COX, D.R., MILLER, H.D.
Stochastic Processes, Chapman & Hall, London, 1972

/30/ ARNDT, K., FRANKEN, P.
Random Point Processes Applied to Availability Analysis of Redundant Systems with Repair, IEEE Trans. Reliability R-22, 1977, pp. 266-269

/31/ JANSEN, U.
Stationäre Verfügbarkeit und Unempfindlichkeit der Zustandswahrscheinlichkeiten - Formeln für die Zuverlässigkeitstheorie, Elektronische Informationsverarbeitung und Kybernetik (Journal of Information Processing and Cybernetics) Vol. 16, No 10/12, Berlin 1980

/32/ NEUTS, M.
Matrix-Geometric Solutions in Stochastic Models, An Algorithmic Approach, Johns Hopkins University Press, Baltimore and London, 1981

/33/ PAZ, A.
Introduction to Probabilistic Automata, Academic Press, New York 1971

/34/ GNEDENKO, B.W., KOWALENKO, I.N.
Einführung in die Bedienungstheorie, R. Oldenbourg, München, 1971

/35/ CORAZZA, M.
Techniques Mathematiques de la Fiabilité Previsionelle des Systèmes (Mathematical Techniques of Reliability Prediction of Systems, in French), Cepadues Editions, Toulouse 1976

/36/ KEMENY, J.G., SNELL, J.L.
Finite Markow Chains, Van Norstrand, New York, 1960

/37/ CINLAR, E.
Introduction to Stochastic Processes, Englewood Cliffs, N.J., 1975

/38/ MEYER, J.F.
On Evaluating the Performability of Degradable Computing Systems in Proc. of FTC 5-8 (June 1978, Toulouse, France) IEEE Computer Society, pp. 44-49

/39/ ESARY, J.D., ZIEHMS, H.
Reliability Analysis of Phased Missions in Reliability and Fault Tree Analysis, p. 213-236, Editors: R.E. Barlow, J.B. Fussel, N.D. Singpurwalla, Society for Ind. Appl. Mathematics, Philadelphia, 1975

/40/ PEDAR, A., SARMA, V.
Phased Mission Analysis for Evaluating the Effectiveness of Aerospace Computer Systems. IEEE-Trans. Rel. R-30, No. 5, p. 429-437, 1981

/41/ SCHRIEFER, D., VOGES, U., WEBER, G.
Design and Construction of a Reliable Microcomputer-Based LMFBR Protection System in Nuclear Power Plant Control and Instrumentation, (Proc. Symp. München 1982), IAEA-SM-265 Vienna 1983, p. 355-366

/42/ WEBER, G.
Failure Diagnosis and Fault Tree Analysis, KfK 3384, July 1982 Kernforschungszentrum Karlsruhe