NASA Contractor Report 178238

ON THE STOCHASTIC DISSEMINATION OF
FAULTS IN AN ADMISSIBLE NETWORK

A. Kyrala

RESEARCH TRIANGLE INSTITUTE
Research Triangle Park, North Carolina

Contract NAS1-17964
January 1987

National Aeronautics and
Space Administration

Langley Research Center
Hampton, Virginia 23665

TABLE OF CONTENTS

ON THE STOCHASTIC DISSEMINATION

OF FAULTS IN AN ADMISSIBLE NETWORK

by A. Kyrala

1. INTRODUCTION

It is intended to discuss the dynamic distribution of faults in a general type of network to be defined in the next section which will be designated as "admissible." The starting point is a UNIQUELY BRANCHED NETWORK in which each pair of nodes is connected by a single branch. Later, the extension to MULTIPLE BRANCHED NETWORKS in which the formerly unique branches are replaced by two or more branches each will be discussed under the subject of REDUNDANCY IN NETWORKS Sec. 11.

The basic discrete model used here is the MARKOV CHAIN, although the extension to a SEMI-MARKOV chain will be discussed later.

2. NETWORK MODIFICATION

It will be supposed that there exists a discrete clock time universal for the entire network with a fundamental time interval τ and that each branch transit time is an integer multiple of τ. In an arbitrary network, this may be approximated by inserting additional new (bipolar) nodes into branches with original transit times larger multiples of τ. If a network (in original form) is such that a signal can be delayed by a multiple of τ at a node, this delay is equivalent to a zero delay at the node followed by insertion of an appropriate number of bipolar nodes into the output branches of the node.

Generally, the actual network to be treated will be unlike the uniquely branched network, which is the starting configuration to be analyzed, however the actual network can be obtained from the uniquely branched one by deletion of branches, node insertions and the addition of redundant parallel branches. The so modified network will be called an ADMISSIBLE NETWORK.

3. MARKOV CHAIN

The general n-node uniquely branched network (which is not necessarily two-dimensional) has a (triangular) number of branches given by

$$\Delta_{n-1} = n(n-1)/2 \qquad (3.1)$$

Let p_{jt} denote the absolute probability that a signal has reached the jth node at epoch[1] t. Supposing that sufficient bipolar nodes have already been inserted so that each time interval τ represents a possible transition period between adjacent nodes, let a_{ijt} denote the conditional probability of transit from node j to node i (i.e., through branch j to i) during the time interval $(t,t+\tau)$ contingent upon the signal having attained node j at epoch t.

The post-transition probability of occupancy of node i at epoch $(t+\tau)$ is then taken to be the linear homogeneous combination of the pre-transition probabilities of occupancy given by the following expression[2]

$$p_i\big|_{(t+\tau)} = \sum_j a_{ijt}\, p_{jt} \qquad (3.2)$$

for i=1 to n subject to the principle of causality (for each epoch t)

$$\sum_j a_{jit} = 1 \tag{3.3}$$

which sets the direction of time and arranges for the transition matrix to have columns summing to unity. It is also supposed that the components of each occupancy vector sum to unity in keeping with its stochastic interpretation.

Multiplication of (3.3) by p_{it} and subtraction from (3.2) then yields

$$p_i\big|_{(t+\tau)} - p_{it} = \sum_j \left(a_{ijt}\, p_{jt} - a_{jit}\, p_{it} \right) \tag{3.4}$$

for i=1 to n giving the change in occupancy probability as a sum of differences between absolute probabilities of transition into and out of i.

One notes in passing that (3.4) exhibits the sufficiency of the Principle of Detailed Balance (with absolute, not conditional probabilities of transition) to ensure stationarity of the Markov chain characterized by the vanishing of the left side of (3.4).

A System for which (3.2) and (3.3) hold is called a MARKOV CHAIN and includes as special cases the Fermi-Dirac and Einstein-Bose statistics, the Diffusion equation, the Boltzmann transport equation as well as (in complex generalization) the Schroedinger and Dirac equations of Quantum Mechanics.[3]

A concrete example of such a Markov chain is afforded by a system which possesses only two states "operative", designated by the subscript o or "inoperative", designated by the subscript i. Suppose that the system undergoes transitions between these states for a very long time. Each transition is characterized by the chain equations

$$p_o^{\,+} = a_{oo}\, p_o + a_{oi}\, p_i \tag{3.5}$$

$$p_i^{\,+} = a_{io}\, p_o + a_{ii}\, p_i \tag{3.6}$$

and the causality conditions,

$$a_{oo} + a_{io} = 1 \tag{3.7}$$

$$a_{oi} + a_{ii} = 1 \tag{3.8}$$

as well as the normalization condition

$$p_o + p_i = 1 \tag{3.9}$$

with the + indicating post-transition absolute probabilities and

p_o = absolute pre-transition probability that system is operative

p_i = " " " " " " inoperative

a_{oo} = conditional probability that system remains operative after the transition CONTINGENT upon having been operative before the transition

a_{oi} = conditional probability that system becomes operative after the transition CONTINGENT upon having been inoperative before the transition

a_{io} = conditional probability that system becomes inoperative after the transition CONTINGENT upon having been operative before the transition

a_{ii} = conditional probability that system remains inoperative after the transition CONTINGENT upon having been inoperative before the transition

Under stationary conditions (after a great many transitions) the + may be removed (i.e., no further change in absolute probabilities occurs) so that

$$p_o = a_{oo}\, p_o + a_{oi}\, p_i \tag{3.10}$$

$$p_i = a_{io}\, p_o + a_{ii}\, p_i \tag{3.11}$$

Using the causality conditions (3.7), (3.8), one then concludes that detailed balance holds for the absolute probabilities of transitions between distinct states. Thus,

$$a_{io}\ p_o = a_{oi}\ p_i \tag{3.12}$$

Hence,

$$p_o/p_i = a_{oi}/a_{io} \tag{3.13}$$

Adding 1 to each side this yields

$$p_i = a_{io}/(a_{io} + a_{oi}) \tag{3.14}$$

and

$$p_o = a_{oi}/(a_{io} + a_{oi}) \tag{3.15}$$

The fraction of transitions during which the system is operative is given by

$$f = N_o/(N_o+N_i) = a_{oi}/(a_{oi} + a_{io}) = p_o \tag{3.16}$$

while the expected number of transitions for recurrence of the inoperative state is given by

$$N(\mathrm{inop}\rightarrow \mathrm{inop}) = 1/p_i = (a_{oi} + a_{io})/a_{io} \tag{3.17}$$

and the expected number of transitions for recurrence of the operative state is given by

$$N(\mathrm{op}\rightarrow \mathrm{op}) = 1/p_o = (a_{oi} + a_{io})/a_{oi} \tag{3.18}$$

The expected number of transitions for the first passage from inoperative to operative state is given by

$$N(\text{inop} \rightarrow \text{op}) = 1/a_{oi} \tag{3.19}$$

based upon the assumption that the contingency of starting inoperative was actually fulfilled. Similarly, the expected number of transitions for first passage from operative to inoperative state is given by

$$N(\text{op} \rightarrow \text{inop}) = 1/a_{io} \tag{3.20}$$

based upon the assumption that the contingency of starting in the operative state was actually fulfilled.

Thus, all of these quantities may be expressed in terms of the conditional probabilities of transition subject to the assumptions stated.

4. FAILURE-RELATED INTERPRETATION OF TRANSITION MATRICES

For a uniquely branched network each off-diagonal element a_{ijt} of the transition matrix corresponds to the traversal of the j to i branch in the specified direction. If a particular branch is deleted, BOTH terms a_{ijt} AND a_{jit} symmetrically located with respect to the main diagonal must be set equal to zero. Also, if the branch connecting the ith and jth nodes fails BIDIRECTIONALLY, the same two terms must be set equal to zero. In a network with UNIDIRECTIONAL branches (say, j to i) only one of the two symmetrants will be non null and this must be set equal to zero. It is in this way that the elements of the transition matrix are related to BRANCH FAILURES.

It is not difficult to construct matricial operators which remove elements from a matrix. Let E_{ii} denote a matrix with a unit element at i,i on the main diagonal and zeros for all other entries. Then for a given transition matrix A, the matrix $E_{ii}\ A\ E_{jj}$ is a matrix in which the element a_{ij} is unaffected, but all other elements are reduced to null. Hence, $A - E_{ii}\ A\ E_{jj}$ is a matrix which has identical elements as A except for a_{ij}, which is replaced by zero. This matrix might reasonably be termed a BRANCH ANNIHILATOR.

The diagonal elements a_{iit} are associated with NODAL DELAY of signal at node i at epoch t. If there is no nodal delay this diagonal element is null at epoch t.

What of the less commonly treated case of NODAL FAILURES? Here it becomes a question of what constitutes a "nodal failure". A given row of the transition matrix (except for the diagonal element) is associated with all INPUTS to the node of the same row number. A given column of the transition matrix is associated (except for the diagonal element) with all OUTPUTS from the node of the same column number. If by NODAL FAILURE is meant (1) failure of all outputs, or (2) failure of all inputs, or (3) failure of all outputs AND all inputs then all off-diagonal elements of the (1) column, or (2) row, or (3) column AND row with the same number as the node must be set equal to zero. In a more elaborate definition of nodal failure, subsets of these entities could be annihilated.

5. VECTOR-MATRICIAL FORMULATION OF THE MARKOV CHAIN

The occupancy probabilities for epoch t may be conceived as components of a STATE VECTOR $\mathbf{p}_t$ while those at epoch $(t+\tau)$ are the components of a state vector $\mathbf{p}_{t+\tau}$ and these two state vectors are related by the transition matrix A_t for epoch t

$$\mathbf{p}_{t+\tau} = \mathbf{A}_t \, \mathbf{p}_t \tag{5.1}$$

All of these occupancy vectors are in the first n-tant since all components of the (n-dimensional) vectors are non-negative. The vector symmetrically directed with respect to the n orthogonal axes has a transpose (row vector) given by

$$\mathbf{u}^T = (1,1,1,\ldots 1,1) \tag{5.2}$$

with all components defined to be unity and T denoting transpose. The normalization of occupancy probabilities then requires that

$$\mathbf{u}^T \cdot \mathbf{p}_t = 1 \tag{5.3}$$

for all epochs t. A state vector of equal likelihood each of whose components is $1/n$ may also be constructed. It should be clearly recognized that the above condition does not ensure that the state vectors retain the same magnitude after transition as they had before transition. Each transition has the potentiality of changing both direction and magnitude of the occupancy state vector $\mathbf{p}_t$ since according to (5.3), it must terminate on a hyperplane orthogonal to the state vector of equal likelihood both before and after transition. The only other restriction is the requirement that the state vectors lie in the first n-tant where all their components will be non-negative. For a large number n of states the "angular separation" θ of a particular state vector from the state vector of equal likelihood is easily estimated to be

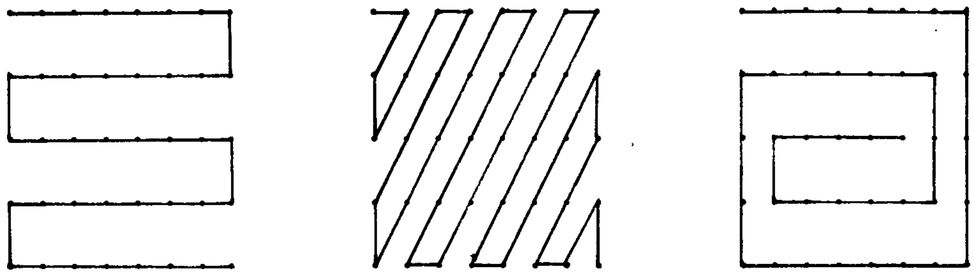
$$\theta = \arccos[(n \sum_k p_k^{\,2})^{-1/2}] \qquad (5.4)$$

The maximum possible angular separation between a state vector and the equal likelihood state vector is given by $\theta = \arccos(1/n^{1/2})$.

6. LINEAR MAPPING OF GRID NETWORK

A GRID NETWORK is a network with nodes at all lattice points of a rectangular lattice with branches vertical or horizontal connecting these points and no others. The occupancy probabilities for the network nodes are the components of the STATE VECTOR in the Markov chain model of the system describing the progress of signal or fault through the network. Therefore, the state of the system is specified in terms of a one-dimensional array of nodes. In terms of sequential occupation of nodes in an actual two-dimensional network, it is more convenient to specify the nodes as a two-dimensional array. Without specifying the geometrical array of nodes, the sequential occupation of states in the Markov model will not have a unique relationship to the occupation of nodes in a given two-dimensional array. This comes about because there does not exist an a priori unique (mapping) correspondence between arrays of different dimensionality.

In particular it is necessary to specify the correspondence between a rectangular grid of nodes at (i,j) with $(i=1,2,...m)$, $(j=1,2,...n)$ and a linear array with $(k=1,2,\ldots,mn)$. Some possible ways of constructing such a correspondence are illustrated in Fig. 1 below:



(a) (b) (c)

Fig. 1 Grid/Linear Array Mappings

The Markov model has no a priori explicit cognizance of the way the two-dimensional array is formed. Generally, it will be most convenient to specify the Markovian sequence of states by the (boustrophedon) path of (a), for which with n columns and m rows the original (one-dim array) Markov nodal number k is given in terms of the (i,j) mapped nodal coordinate by

$$k = ni - (-1)^i j - (n-1)/2 + (-1)^i (n+1)/2 \qquad (6.1)$$

for $(i=1,2\ldots m)$ and $(j=1,2\ldots n)$. The inverse mapping yielding (i,j) for a given value of k is found as follows. The row number i is given by

$$i = [(k-1)/n] + 1 \qquad (6.2)$$

where [] means "integer part of". Then j is given by

$$j = (-1)^{i}(ni - k - (n-1)/2) + (n + 1)/2 \tag{6.3}$$

In this way (with such a specified path) the nodal occupancy probabilities p_k may be replaced by p_{ij} where the location of the node (i,j) is specified in a two-dimensional grid. If p_{ijt} is the a priori absolute probability of occupancy of the node at (i,j) at epoch t, then one can introduce

${}_{uv}a_{ijt}$ = conditional probability of transit from (i,j) to (u,v) during the time interval $(t,t+\tau)$

The Markov chain equation then becomes

$$p_{uv\ t+\tau} = \sum_{i,j} {}_{uv}a_{ijt}\ p_{ij} \tag{6.4}$$

with

$$1 = \sum_{i,j} p_{ijt} \tag{6.5}$$

for occupancy normalization and

$$1 = \sum_{u,v} {}_{uv}a_{ijt} \tag{6.6}$$

as causality principle.

The same equations can be more concisely represented by introducing the GAUSSIAN (complex) integers defined by

$g = i + j\sqrt{-1} \qquad 1<i<m\ ,\ 1<j<n$

$f = i' + j'\sqrt{-1} \qquad 1<i'<m,\ 1<j'<n$ Then one has

$$p_{f\,t+\tau} = \sum_g f^a{}_{gt} \; p_{gt} \tag{6.7}$$

with occupancy normalization

$$1 = \sum_g p_{gt} \tag{6.8}$$

and causality principle

$$1 = \sum_f f^a{}_{gt} \tag{6.9}$$

with all quantities having complex indices.

If it is desired to restrict to "nearest neighbor transitions" so that (i,j) to $(i+1,j)$, $(i-1,j)$, $(i,j+1)$, $(i,j-1)$ are the only transitions from (i,j) with non-null conditional probabilities of transition one can find these transitions in terms of the original k sequence. Thus, the transitions considered in terms of k become

$$(i+1,j):\ k = n(i+1) + (-1)^i j - (n-1)/2 - (-1)^i (n+1)/2 \tag{6.10}$$

$$(i-1,j):\ k = n(i-1) + (-1)^i j - (n-1)/2 - (-1)^i (n+1)/2 \tag{6.11}$$

$$(i,j+1):\ k = ni - (-1)^i (j+1) - (n-1)/2 + (-1)^i (n+1)/2 \tag{6.12}$$

$$(i,j-1):\ k = ni - (-1)^i (j-1) - (n-1)/2 + (-1)^i (n+1)/2 \tag{6.13}$$

for the two-dimensional post-transition states indicated.

7. THE DIFFUSION AND PROPAGATION OF FAULTS OR SIGNALS IN A NETWORK

It will now be shown under what conditions it is possible to have a diffusive or wavelike propagation of successive faults or signals in a network. In order to get such a propagation in a Markov model, it is necessary to impose a SELECTION RULE restricting to transitions between nearest neighbors and it is important to distinguish between chains obeying (3.2), (3.3) or (6.6), (6.8). Both are Markov models but the nearest neighbors are different in each. Propagation in the chain of Section 3 means propagation through a linear array of states while propagation in the chain of Section 6 means propagation through a two-dimensionally ordered set of states. Generally, it is not possible to get a wavelike propagation through the states in either case without imposing some restrictions on the transition probabilities of the general Markov chains of either section.

For the case of a transition matrix independent of time the conditions for wave-like propagation can be readily adduced.
The chain equation (3.2) and the causality principle (3.3) by imposition of the selection rule[5]

$$|i - j| > 1 \text{ implies } a_{ij} = 0 \qquad (7.1)$$

become

$$p_i|_{t+\tau} = a_i|_{i+1}\, p_{i+1}|_t + a_i|_{i-1}\, p_{i-1}|_t + a_{ii}\, p_{it} \qquad (7.2)$$

and

$$a_{ii} + a_{i+1}|_i + a_{i-1}|_i = 1 \qquad (7.3)$$

The selection rule simply excludes transitions except among nearest neighbors. If h is the mean number of states through which a fault propagates during transition time τ, one can then define the quantities D_i, w_i, μ_i as follows:

$$2\tau D_i = h^2 \left(a_i|_{i+1} + a_i|_{i-1} \right) \tag{7.4}$$

$$\tau w_i = h \left(a_i|_{i+1} - a_i|_{i-1} \right) \tag{7.5}$$

$$\tau\mu_i = \left(a_{i+1}|_i - a_i|_{i+1} \right) + \left(a_{i-1}|_i - a_i|_{i-1} \right) \tag{7.6}$$

Using the so defined quantities and (7.3), the restricted chain equation (7.2) can be written in the form

$$(p_i|_{t+\tau}-p_{it})/\tau = D_i(p_{i+1}|_t-2p_{it}+p_{i-1}|_t)/h^2+w_i(p_{i+1}|_t-p_{i-1}|_t)/2h - \mu_i p_{it} \tag{7.7}$$

which is a finite approximant of the diffusion equation with drift w and rate of destruction μ (supposing $(a_{i+1}|_i + a_{i-1}|_i)>(a_i|_{i+1} + a_i|_{i-1})$)

$$\partial_t p = D\,\partial_x{}^2 p + w\,\partial_x p - \mu p \tag{7.8}$$

with diffusion coefficient D. This indicates that with transitions restricted to nearest neighbors faults may diffuse through the states.

μ may also be replaced by $-\mu$ provided only that $(a_i|_{i+1} + a_i|_{i-1})>(a_{i+1}|_i + a_{i-1}|_i)$. Thus the expression given for μ functions as a fault annihilator or creator.

For the case where

$$h \ll (a_i|_{i+1} - a_i|_{i-1})/(a_i|_{i+1} + a_i|_{i-1}) \tag{7.9}$$

the diffusive term will become negligible with respect to the drift term and the fault will propagate through the network in a wave-like fashion provided only that μ_i is zero. The phase velocity is naturally w_i.

If the individual states are arranged in a two-dimensional array rather than the one-dimensional linear array above, a regrouping of conditional probabilities to give a "diffusive case" or possibly a wave-like case can still be attained by imposing a selection rule on the transitions. However, it is most important to realize that in such a development the definition of "nearest neighbors" has changed and the analysis must take this into account.

For the case where the transition matrix is a function of time, it is more convenient to return to the vector-matricial model of Section 5. Put the case that at some time the state vector from some index on has only null components (unoccupied states). The question is then posed as to what conditions the transition matrix must fulfill in order to advance the occupancy state by contiguous state as the transitions occur. If the index (component number) from which all previous components are not necessarily zero is q, then the qth and all later components are taken to be zero. In order that the transition matrix now accomplish the extension of occupancy to qth component of the state vector BUT NOT BEYOND it will be sufficient if all elements of the transition matrix with row numbers greater than q and column numbers less than q be null. Thus, it is readily grasped that not only is a wave of replacement of zeros propagating in the state vector but also a wave of zero replacements is simultaneously occurring in the transition matrix. Hence, it is seen that with the fulfillment of these conditions faults can propagate in a wave-like fashion, even in the case where the transition matrix is time dependent. An example of such a propagation is given below with the convention that 1 does not represent the unit but rather any non-null element. Then schematically one has

$$\begin{bmatrix} 1&1&1&1&1&1&1&1 \\ 1&1&1&1&1&1&1&1 \\ 0&1&1&1&1&1&1&1 \\ 0&1&1&1&1&1&1&1 \\ 0&1&1&1&1&1&1&1 \\ 0&1&1&1&1&1&1&1 \\ 0&1&1&1&1&1&1&1 \\ 0&1&1&1&1&1&1&1 \end{bmatrix} \begin{vmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{vmatrix} = \begin{vmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{vmatrix}$$

(a)

$$\begin{bmatrix} 1&1&1&1&1&1&1&1 \\ 1&1&1&1&1&1&1&1 \\ 1&1&1&1&1&1&1&1 \\ 0&0&1&1&1&1&1&1 \\ 0&0&1&1&1&1&1&1 \\ 0&0&1&1&1&1&1&1 \\ 0&0&1&1&1&1&1&1 \\ 0&0&1&1&1&1&1&1 \end{bmatrix} \begin{vmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{vmatrix} = \begin{vmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{vmatrix}$$

(b)

$$\begin{bmatrix} 1&1&1&1&1&1&1&1 \\ 1&1&1&1&1&1&1&1 \\ 1&1&1&1&1&1&1&1 \\ 1&1&1&1&1&1&1&1 \\ 0&0&0&1&1&1&1&1 \\ 0&0&0&1&1&1&1&1 \\ 0&0&0&1&1&1&1&1 \\ 0&0&0&1&1&1&1&1 \end{bmatrix} \begin{vmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{vmatrix} = \begin{vmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{vmatrix}$$

(c)

$$\begin{bmatrix} 1&1&1&1&1&1&1&1 \\ 1&1&1&1&1&1&1&1 \\ 1&1&1&1&1&1&1&1 \\ 1&1&1&1&1&1&1&1 \\ 1&1&1&1&1&1&1&1 \\ 0&0&0&0&1&1&1&1 \\ 0&0&0&0&1&1&1&1 \\ 0&0&0&0&1&1&1&1 \end{bmatrix} \begin{vmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{vmatrix} = \begin{vmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{vmatrix}$$

(d)

$$\begin{bmatrix} 1&1&1&1&1&1&1&1 \\ 1&1&1&1&1&1&1&1 \\ 1&1&1&1&1&1&1&1 \\ 1&1&1&1&1&1&1&1 \\ 1&1&1&1&1&1&1&1 \\ 1&1&1&1&1&1&1&1 \\ 0&0&0&0&0&1&1&1 \\ 0&0&0&0&0&1&1&1 \end{bmatrix} \begin{vmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{vmatrix} = \begin{vmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 0 \\ 0 \end{vmatrix}$$

(e)

$$\begin{bmatrix} 1&1&1&1&1&1&1&1 \\ 1&1&1&1&1&1&1&1 \\ 1&1&1&1&1&1&1&1 \\ 1&1&1&1&1&1&1&1 \\ 1&1&1&1&1&1&1&1 \\ 1&1&1&1&1&1&1&1 \\ 1&1&1&1&1&1&1&1 \\ 0&0&0&0&0&0&1&1 \end{bmatrix} \begin{vmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 0 \\ 0 \end{vmatrix} = \begin{vmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 0 \end{vmatrix}$$

(f)

Fig. 2

which indicates in a graphic way what is meant by "propagation" through the transition matrix simultaneous with the propagation through the states of the state vector. The propagation is that of a partition between null and non-null states and null and non-null transitions. Fig. 2 then corresponds to the propagation of nulls in the state vector with $p_k=0$ for $k>m-1$ and in the transitions matrix with $a_{jk}=0$ for $j>m>k$.

From this it is seen that the transitions must have a very particular type of time dependence (inhomogeneity) in order for propagation as such to occur.

Regardless of whether it occurs or not, one can form useful estimates of the concentrative or dispersive effect of each transition by calculating the expected state and expected standard deviation in states after each transition. Thus,

$$\langle k \rangle = \sum_k k\, p_{kt} \tag{7.10}$$

$$\sigma_k{}^2 = \sum_k k^2\, p_{kt} - \langle k \rangle^2 \tag{7.11}$$

both of which are quite naturally time dependent. From the view point of the transition matrix to effect a concentration of the occupancies in the state vector on any particular transition the row vectors which form the transition matrix must be close to orthogonal to the state vector on which they operate except in a narrow range of row numbers (in the extreme case 1). On the contrary, if the transition matrix is to effect an equalization of the components of the state vector then the row vectors should all have the same scalar product with the state vector on which they operate. In either case (7.10) and (7.11) describe quantitatively the distribution of occupancy in states. Another measure of how uniformly (or non-uniformly) states are distributed in the state vector is the Entropy defined by $-\Sigma p_k \ln p_k$ which takes the value $\ln n$ for the state vector of equal likelihood and the value 0 for the state vector of a system stochastically certain to be in a particular state.

8. SEMI-MARKOV GENERALIZATION OF MARKOV CHAINS

It has been pointed out in Section 2 that delay of faults and signals could under certain circumstances be treated by nodal insertions in the context of a Markov model. This requires delay times which are multiples of a common (constant) transition time. There is another method which is suited to continuous stochastically variable delay times. This is the method of Semi-Markov Chains. They are constructed around an "embedded" Markov Chain which may be taken to have a time independent transition matrix.

Tau is now taken to be a continuous stochastic transition time and the following definitions apply:

a_{ij} = conditional probability of transition from j to i CONTINGENT upon the system having been in j (i.e., upon j having been occupied before the transition).

$F_{ij}(\tau)$ = conditional probability of transition from j to i in a time interval less than τ CONTINGENT upon the transition from j to i having occurred.

$a_i(\tau)$ = conditional probability of node i being occupied in time interval less than τ CONTINGENT upon a transition from some node to i having occurred.

With these definitions we obtain from them the SEMI-MARKOV Chain equation

$$a_i(\tau) = \sum_j F_{ij}(\tau)\, a_{ij} \tag{8.1}$$

and the principle of causality

$$1 = \sum_i a_{ij} \qquad \text{with } a_{ij} > 0 \tag{8.2}$$

If τ is allowed to become infinite then F_{ij} and a_i should both become unity. However,

$$F_{ij}(\infty) = 1 \text{ implies } a_i(\infty) = 1 \text{ only if } \sum_j a_{ij} = 1 \tag{8.3}$$

Hence, the transition matrix must be doubly stochastic. It is parenthetically noted that since reconfigurations which "restore" the condition of the network in some sense are being considered, this may very well be appropriate for the cases at hand. In any case

$$F_{ij}(0) = 0 \quad \text{implies} \quad a_i(0) = 0 \tag{8.4}$$

so that no instantaneous transitions are allowed.

The (Stieltjes) differential of both sides of (8.1) is then

$$da_i(\tau) = \sum_j a_{ij}\, dF_{ij}(\tau) \tag{8.5}$$

and both differentials are clearly non-negative. The normalization consistent with (8.3) is

$$\int_0^\infty da_i(\tau) = 1 = \int_0^\infty dF_{ij}(\tau) \tag{8.6}$$

so that the mean time interval to occupy the ith node (during the transition) is

$$\tau_i = \int_0^\infty \tau \ da_i(\tau) \tag{8.7}$$

and the mean transition time for the j to i (nodal) transition is defined to be

$$\tau_{ij} = \int_0^\infty \tau \ dF_{ij}(\tau) \tag{8.8}$$

so that the conclusion

$$\tau_i = \sum_j a_{ij} \tau_{ij} \tag{8.9}$$

implies that the mean time interval required to occupy node i is a weighted average of the mean transition times into the node which is a consequence of the double stochasticity of the (embedded) transition matrix. The mean time interval to occupy all n states of the Semi-Markov chain is given by

$$\langle\tau\rangle = (1/n) \sum \tau_i \tag{8.10}$$

It should be noted by the reader that the entire formulation of the Semi-Markov chain is in terms of conditional probabilities. If it were desired to generalize the chain equations (3.2) involving absolute probabilities of occupancy one should have

$$p_i\big|_{t+\tau} = \sum_j F_{ij}(\tau)\, a_{ij}\, p_{jt} \tag{8.11}$$

with a clear understanding of the difference between universal clock time (epoch) t and stochastic transition time interval τ. The difficulty with such an extension is that even if the occupancy probabilities are initially referenced to clock time they become functions of the stochastic transition times after any transitions introducing numerous new variables into the problems.

9. STATISTICAL INDEPENDENCE OF STATES

It is a tacit assumption of the Markov chain concept that the states must be defined so that they can be occupied independently and the same requirement applies in principle to the semi-Markov chain which contains an embedded Markov as part of its structure. In the semi-Markov chain the situation is even more severe with a sparse transition matrix because the consistent calculation of mean transition times requires that the transition matrix be doubly stochastic. The models[6] of CARE III, SURE, HARP, etc., seem to overlook this fact and are therefore dealing with state definitions which are NOT INDEPENDENT. Because of this they should not be referred to as semi-Markov systems. This remark does not of itself invalidate the calculation of path transit probabilities made in those systems either in the time domain[7] or in the frequency domain[8].

10. COMMENTS ON VOTER SYSTEMS

The voter system of n elements yields "agreement" for k failures among the n elements provided $n-k>[n/2]$ ([] means integer part of). Otherwise the voter system yields "disagreement". It is a majority rule system.

The individual voter elements are however subject to malfunction hence the choice between operative and inoperative for the system as a whole can occasionally occur without reference to input. This would be the case of the "irrational voter" whose choices are entirely random. The probability of agreement[9] on such a random basis is

$$p_a = \sum_{k=0}^{[n/2]} C^n_k \, p^{n-k} \, q^k \qquad (10.1)$$

where C^n_k is the (binomial coefficient) number of combinations of n things taken k at a time, p is the probability of a YES vote by an individual element, q is the probability of a NO vote by an individual element. Thus p_a might be called the "probability of irrational agreement" (e.g., an agreement to go to war when it serves no known national interest). In terms of the expected number of agreements $N_a = 1/p_a$, expected number of YES votes $N_Y = 1/p$ and expected number of NO votes $N_N = 1/q$; one has from (10.1).

$$N_a = 1/(\Sigma_{k=o} \, C^n_k \, /(N_Y^{n-k} \, N_N^k)) \qquad (10.2)$$

The probability of agreement based upon rational factors is undoubtedly not binomial. Since the elements of the voter system are superficially identical, it seems they could be reasonably assumed to be equicorrelated because of their common function but hardly independent. Their common design could apparently yield a correlative bias in performance. Thus, from the total expected number of agreements of the voter system should be subtracted the expected number of irrational agreements given by (10.2) to arrive at the expected number of rational agreements (i.e., agreements arrived at solely by mutual consideration of inputs). In future work modeling the correlation between individual voter elements should be of considerable importance.

11. REDUNDANCY IN NETWORKS

The principal device used to increase reliability of networks is BRANCH REDUNDANCY in which a branch with probability q of failure by itself is replaced by n parallel branches with probability of failure q^n (on the assumption that the parallel branches fail independently). From this it can be readily calculated that the number of branches required to reduce the probability of the failure of the multi-branched system to 10^{-m} is simply

$$n = [\ m/\log_{10}(1/q)\ ] \qquad (11.1)$$

from which a small table may be constructed with the values of n in the body of the table and the values of q as vertically arrayed entries and the values of m as horizontally arrayed entries

q				
.1	6	9	12	
.01	3	5	6	
.001	2	3	4	
	6	9	12	m

Table 3

Branches with a greater probability of failure are also easily calculated from (11.1).

The calculation of failure probability of parallel multibranches is accomplished by successive application of the calculation for two branches say

1 and 2 in parallel. Then the probability of failure of the double branch is simply $q_1 q_2$ where the "q"s are the probabilities of failure of individual branches.

At this stage it is easy to determine the effect on the transition matrix of the uniquely branched network. Corresponding to the branch a_{ij} this conditional probability must be replaced by the probability of the multi-branch.

The question of NODAL REDUNDANCY would seem to imply replacing a single node by n nodes but this cannot be done without simultaneously multiplying all inputs and outputs for the node which considerably complicates the network. Apparently the use of a voter system is another way of handling the nodal redundancy problem. In that case the node complete with its treatment of inputs is replaced by a "new kind of node" capable of making its own decisions about how to treat inputs.

12. NON-STATIONARY FAULT ARRIVAL RATE THEORY

In view of the importance of fault arrival rates it seems worthwhile to attempt to construct a theory to handle this parameter under non-stationary conditions. As a first approximation this will be based upon two assumptions:

Assumption 1: The ratio R of reconfiguration rate to fault arrival rate u is a constant.

Assumption 2: The ratio ε of the absolute probability of the transition from operative system state to inoperative system state to the absolute probability of the transition from inoperative system state to operative system state is a constant.

Two ways of the system becoming inoperative contingent upon its having been operative will be recognized. The system may become inoperative due to internal malfunction quite independently of fault arrival or it may become inoperative due to fault arrival. In the latter case it will be reasonable to expect the effect to be proportional to the fault arrival rate u. The following definitions apply:

$b_{io}\, u$ = conditional probability of system becoming inoperative due to fault arrival contingent upon having been operative

c_{io} = conditional probability of system spontaneously becoming inoperative contingent upon having been operative

c_{oi} = conditional probability of system spontaneously becoming operative contingent upon having been inoperative

$b_{oi}\, Ru$ = conditional probability of system becoming operative (due to reconfiguration capability) contingent upon having been inoperative

$a_{oi}\, p_i$ = absolute probability of system becoming operative from inoperative

$a_{io}\, p_o$ = " " " " " inoperative from operative

The term "astationarity parameter" will be used for ε. Only for $\varepsilon = 1$ does stationarity obtain. The principle of astationarity

$$a_{io}\, p_o = \varepsilon\, a_{oi}\, p_i \tag{12.1}$$

now replaces the stationarity condition. The conditional probabilities of transition may now be expressed in terms of the definitions above

$$a_{io} = b_{io}\, u + c_{io} \tag{12.2}$$

$$a_{oi} = b_{oi}\, Ru + c_{oi} \tag{12.3}$$

substituting these into (12.1) yields

$$(b_{io}\, u + c_{io})\, p_o = \varepsilon\, (b_{oi}\, Ru + c_{oi})\, p_i \tag{12.4}$$

Solving this for u then yields

$$u = (\varepsilon\, c_{oi}\, p_i - c_{io}\, p_o)/(b_{io}\, p_o - \varepsilon R\, b_{oi}\, p_i) \tag{12.5}$$

If now an Ansatz such as

$$p_i = \int_o^t dF(t) \tag{12.6}$$

$$p_o = \int_t^\infty dF(t) \tag{12.7}$$

with

$$\int_o^\infty dF(t) = 1 \tag{12.8}$$

is used so that the probability of being inoperative initially is taken to be zero as is the probability of being operative ultimately. It should be noted that $F(t)$ is not a distribution function or the occupancy probabilities would be constrained to be monotonic. In any case the fault arrival rate becomes

$$u = \frac{\varepsilon\, c_{oi} \int_o^t dF(t) - c_{io} \int_t^\infty dF(t)}{b_{io} \int_t^\infty dF(t) - \varepsilon\, R\, b_{oi} \int_o^t dF(t)} \tag{12.9}$$

With more special assumptions about the occupancy probabilities other forms of (12.5) become possible. If it were assumed that the relative probability of being inoperative to that of being operative became exponentially unlikely with increasing time (12.5) would become

$$u = \varepsilon K t / (b_{io} e^{\lambda t} - \varepsilon R b_{oi}) \tag{12.10}$$

under the assumptions $c_{io} = 0$ and $p_o = 1/(1 + e^{-\lambda t})$ and $p_i = e^{-\lambda t}/(1 + e^{-\lambda t})$ so that $p_i/p_o = e^{-\lambda t}$ and $c_{oi} = Kt$ which may be fitted to data if the coefficients are constant.

Finally some remarks about fault arrival should be made. In hardware faults don't arrive at failure states, they arrive at devices. In software faults don't arrive at failure states, they arrive at nodes in flow charts.

13. TRANSITION MATRIX CHARACTERIZATION FOR SOFTWARE ERRORS

The principle problem of reliability for software appears to be the masking of errors concealed in a node of the flow chart which is not invoked during a particular sequence of runs. The basic requirement is then a way of comparing the system performance with utilization of this node versus the system performance in the avoidance of this node. As far as the transition matrix is concerned, removal of this node is equivalent to removing the row and column containing the node from the original transition matrix. Then using the two transition matrices one would calculate the probabilities of attaining the same end states (final instructions) for each of the matrices. The ratio of these probabilities would then yield a measure of the potential damage to the program in terms of relative performance times.

NOTES

1. The Riordan convention of calling a particular instant in time an epoch to distinguish it from a time interval will be followed here. See W. Feller: An Introduction to Probability Theory and its Applications, Vol. I, Wiley, NY, which also discusses discrete Markov chains.

2. The vertical line in the subscript emphasizes the separation of two distinct variables in a subscript.

3. A. Kyrala: Selection Rules, Causality and Unitarity in Statistical and Quantum Physics: Foundations of Physics, Vol. 4, No. 1, March 1974, p. 31-51.

5. The large arrow means "implies".

6. See Appendix A.

7. A. L. White: Upper and Lower Bounds for Semi-Markov Reliability Models of Reconfigurable Systems: NASA Contractor Report 172340, April 1984.
A. L. White: Synthetic Bounds for Semi-Markov Reliability Models: NASA Contractor Report 178008.

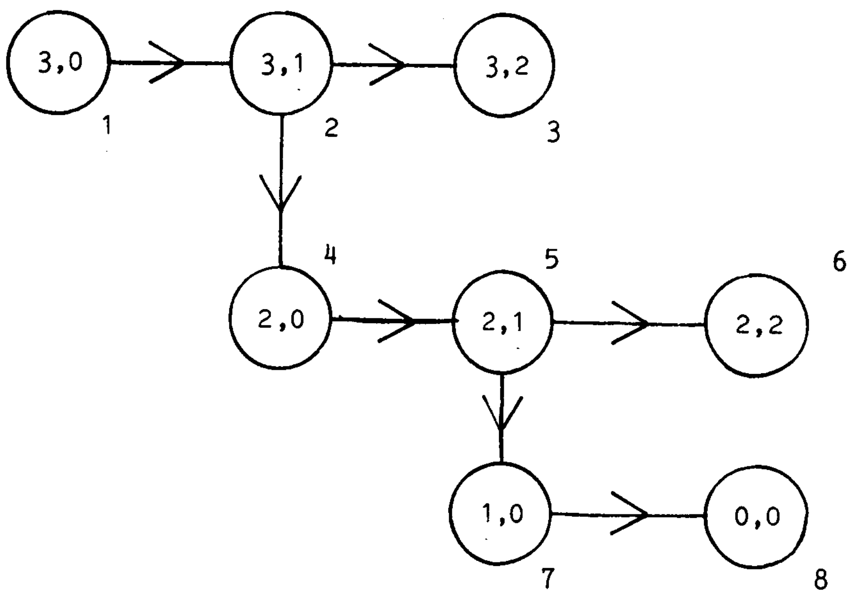
8. See Appendix B.

9. Considering only YES agreements. There is a similar expression for NO agreements.

APPENDIX A

ILLUSTRATION OF TRANSITION MATRIX FOR SURE STATES

Corresponding to the SURE State Diagram shown below



Relabeling the states from pairs of digits (the first being the number of voter elements corresponding to YES, the second being the number of voter elements corresponding to NO) to single digits indicated on the diagram one may construct the transition matrix as follows

$$
\begin{array}{lccccccccc}
 & & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \quad \text{initial states} \\
 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & 2 & a_{21} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\text{final} & 3 & 0 & a_{32} & 1 & 0 & 0 & 0 & 0 & 0 \\
 & 4 & 0 & a_{42} & 0 & 0 & 0 & 0 & 0 & 0 \\
\text{states} & 5 & 0 & 0 & 0 & a_{54} & 0 & 0 & 0 & 0 \\
 & 6 & 0 & 0 & 0 & 0 & a_{65} & 1 & 0 & 0 \\
 & 7 & 0 & 0 & 0 & 0 & a_{75} & 0 & 0 & 0 \\
 & 8 & 0 & 0 & 0 & 0 & 0 & 0 & a_{87} & 1
\end{array}
$$

from which it can be readily discerned that the matrix is too sparse to fulfill the normalizations on rows and columns of the Semi-Markov chain, although the columns can sum to unity satisfying the causality condition of the Markov chain in the case where the transition times become constant.

APPENDIX B

NOTE ON
SEQUENTIAL PATH FAILURE PROBABILITIES
by LAPLACE STIELTJES TRANSFORM

by A. Kyrala

In considering the transmission of signals or faults through a path consisting of bipolar subsections, it is well known that the output of any section is the convolution of the input to that section and the system function for the section. For a linear array of such sections the overall output will be given by a repeated convolution. For four filters in series one has

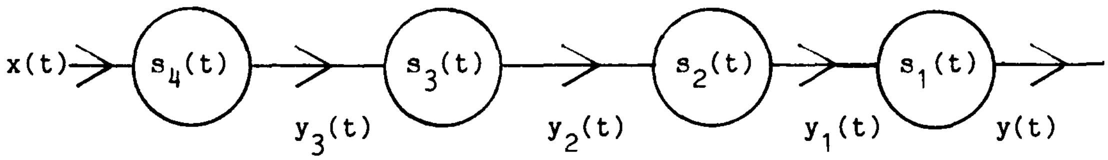


Fig.1

so that the successive convolutions are

$$y_3(t) = \int_0^t s_4(\tau_4)\, x(t-\tau_4)\, d\tau_4 \tag{1}$$

$$y_2(t) = \int_0^t s_3(\tau_3)\, y_3(t-\tau_3)\, d\tau_3 \tag{2}$$

$$y_1(t) = \int_0^t s_2(\tau_2)\, y_2(t-\tau_2)\, d\tau_2 \tag{3}$$

$$y(t) = \int_0^t s_1(\tau_1)\, y_1(t-\tau_1)\, d\tau_1 \tag{4}$$

which upon successive substitutions yields a four-fold multiple integral

$$y(t) = \int_0^{t} \int_0^{t-\tau_1} \int_0^{t-\tau_1-\tau_2} \int_0^{t-\tau_1-\tau_2-\tau_3} s_1(\tau_1)s_2(\tau_2)s_3(\tau_3)s_4(\tau_4)\; x(t-\tau_1-\tau_2-\tau_3-\tau_4)\; d\tau_1 d\tau_2 d\tau_3 d\tau_4 \tag{6}$$

Instead of dealing with (6) as an expression from which output can be calculated one can use the Laplace-Stieltjes transform defined by

$$Y(s) = \int_o^{\infty} e^{-st}\, dF_Y(t) \tag{7}$$

where $F_Y(t)$ is the distribution function for $y(t)$. Using a similar notation for the other elements in Fig. 1 the transformed version of (1), (2), (3), and (4) become

$$Y(s) = S_1(s)\, Y_1(s) \tag{8}$$

$$Y_1(s) = S_2(s)\, Y_2(s) \tag{9}$$

$$Y_2(s) = S_3(s)\, Y_3(s) \tag{10}$$

$$Y_3(s) = S_4(s)\, X(s) \tag{11}$$

Thus instead of (6) one arrives at the transform of the output simply by multiplication

$$Y(s) = \prod_{k=1}^{4} S_k(s)\, X(s) \tag{12}$$

In a similar way any number of elements in series can be treated.

To determine the moments of output (or any intermediate stage), one simply differentiates (7) with respect to s and then lets s approach zero. Thus

$$Y^{(n)}(0) = (-1)^n \int_0^\infty t^n \, dF_Y(t) \tag{13}$$

so that the MacLaurin series for $Y(s)$ is then

$$Y(s) = \sum_{n=0}^\infty Y^{(n)}(0) \, s^n/n! = \sum_{n=0}^\infty (-1)^n \langle t^n \rangle s^n/n! \tag{14}$$

In particular

$$\langle t \rangle = \int_0^\infty t \, dF_Y(t) = - Y'(0) \tag{15}$$

is the mean for the output and the standard deviation σ_t is given by

$$\sigma_t^2 = Y''(0) - [Y'(0)]^2 \tag{16}$$

It should be clearly understood that the elements $s_k(t)$, which are taken to be system functions in filter theory can in the present stochastic context be regarded as failure probability densities associated with subsections of the path.

Standard Bibliographic Page

1. Report No. NASA CR-178238	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle On the Stochastic Dissemination of Faults in an Admissible Network		5. Report Date January 1987
		6. Performing Organization Code
7. Author(s) A. Kyrala		8. Performing Organization Report No.
9. Performing Organization Name and Address Research Triangle Institute P. O. Box 12194 Research Triangle Park, NC 27709		10. Work Unit No.
		11. Contract or Grant No. NAS1-17964
12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Washington, DC 20546		13. Type of Report and Period Covered Contractor Report
		14. Sponsoring Agency Code 505-66-21-01
15. Supplementary Notes Langley Technical Monitor: A. O. Lupton		

16. Abstract

The dynamic distribution of faults in a general type network is discussed. The starting point is a UNIQUELY BRANCHED NETWORK in which each pair of nodes are connected by a single branch. Mathematical expressions for the uniquely branched network transition matrix are derived to show that sufficient stationarity exists to ensure the validity of the use of the Markov Chain model to analyze networks. In addition the conditions for the use of Semi-Markov models are discussed. General mathematical expressions are derived in an examination of branch redundancy techniques commonly used to increase reliability.

17. Key Words (Suggested by Authors(s)) Branched Networks Markov Models Reliability Fault Analysis		18. Distribution Statement Unclassified - Unlimited Subject Category 65	
19. Security Classif.(of this report) Unclassified	20. Security Classif.(of this page) Unclassified	21. No. of Pages 35	22. Price A03

For sale by the National Technical Information Service, Springfield, Virginia 22161

NASA Langley Form 63 (June 1985)