SOLUTION OF INTERACTING SEMI-MARKOV SUBPROCESSES

Jon A. Sjogren

Faculty of Graduate Studies, Towson University

15 July 2015

Abstract

Semi-Markov processes are increasingly being used in the modeling of complex reconfigurable systems (fault-tolerant computers). The estimation of system reliability reduces to solving the model for its state probabilities. When decomposition into smaller submodels is possible, solving the decomposed models and joining the solutions has several advantages, including lower computational cost and more stable numerical properties. Interesting cases arise when non-failure states of the subsystems can be system failure modes. A technique is proposed to combine information about the submodels to estimate this additional probability. This question is considered in the framework of Semi-Markov processes whose transition densities have a *rational* Laplace transform. We state a theorem that characterizes in a simple fashion when these transforms exist, and prove as a special case that the transforms always exist when no "instantaneous jumps" are present. When there are no instantaneous jumps present (distributions and densities are "exponomial functions") the state probabilities are exponomial as well. This gives a way to build up a system hierarchically from smaller systems. The technique is applied to a Semi-Markovian reconfigurable voted sensor submodel and a constant-rate Markovian network submodel. The solution is obtained by our formulas employing a symbolic algebra package.

Introduction

Modern reliability modeling practice involves several techniques, including Fault-Tree analysis and (semi)-Markov processes. Fault-Trees enable one to evaluate the impact of certain dependencies within the entire system. For instance, the fact that the failure of a certain component results in the failure of higher-level component (sub-system) may be modeled. Markov and semi-Markov processes (*chains*) can depict more general types of dependency. For example, consider a triplex system consisting of 3 identical components A_1, A_2, A_3, each with failure rate *lambda*. Then system failure can be represented within the domain of Fault-Trees by a '2 out of 3 gate' as in Figure Int-1a, or by a set of AND and OR gates (Figure Int-1b). Failure may also be depicted by a Markov process (Figure Int-2). With the rates indicated, the probability of failure at time t is the probability that the process is in state F at t.

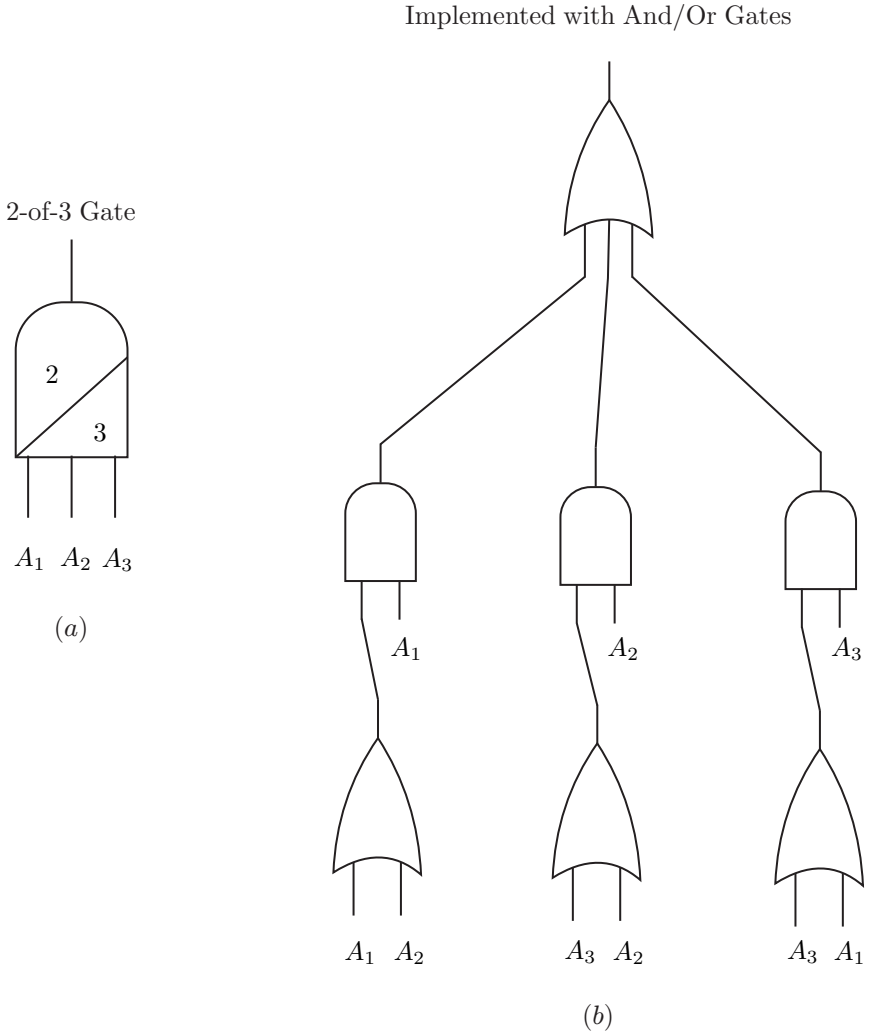


Figure Int-1

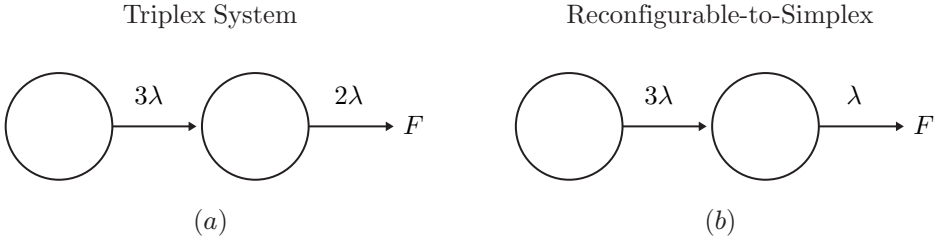


Figure Int-2

Another "fault-tolerant architecture" is a triplex which operates by (instantly) detecting a first fault and then reconfiguring to a simplex system. This is accomplished by unplugging the defective component and a randomly selected "good" component. There is no way, using only the three components in a logic gate (Fault-Tree) arrangement, to represent the event of system failure. But it is easy

to give the corresponding Markov process (Figure Int-2b). Thus the necessity for Markov models tends to come about when *system reconfiguration* is a characteristic feature.

For mission-critical systems found in process-control, avionics, and so on, one is concerned with the reliability at some particular time (mission time). Given a Markov process that models a system, the unreliability or *failure probability* can be found by using a numerical differential equation solver [Reibman & Trivedi]. Alternatively, one may wish to have this quantity in *closed form* as a function of time.

Large models (with many states and transitions) arise naturally in the study of complex reconfigurable systems. The solution of such a model can be both expensive and time-consuming. However, when the model can be *decomposed hierarchically* into smaller models, the process of solving the smaller models is generally less expensive than solving the large one. The desired reliability number associated with the large number can be computed from quantities derived in the solution of the smaller ones.

The major purpose of this article is to increase our understanding of hierarchical modeling and to increase our capabilities for solving such models. Two equations, presented as (4.7) and (4.8), govern the method of solution by decomposition.

In order to write down these decomposition relationships we must present the *Chapman-Kolmogorov equations* somewhat differently than in previous literature. However, [Ross] gives a related treatment. These C-K equations involve quantities (probability density functions) which are used in a new integral equation, (4.7). The solution of the integral equation is a function that expresses part of the failure probability present in a "combined model" that results from two smaller models. The failure mode involved here (in the combined model) does not generally arise from the failure modes of the smaller models. Thus we have developed a technique for combining models accurately, even though there may be some interaction between them, leading to new failure modes.

The obvious advantage of the closed-form framework is that once the solution is found (presumably at significant computational cost), the reliability is easily calculated for any desired value of time t. In addition, it is easier to find sensitivity functions (with respect to failure rates or other parameters). The Symbolic Hierarchical Automated Reliability and Performance Evaluator (SHARPE) program takes this closed-form approach [Sahner & Trivedi1]. The reliability functions encountered in SHARPE are the so-called "exponomial" functions and can be easily represented "symbolically". They consist of probability distribution functions of a particular algebraic form to be described shortly.

The SHARPE modeling framework is amenable to the use of hierarchical modeling techniques. SHARPE was intended to promote hierarchical decomposition. Much of the information needed in applying our decomposition method can be obtained from SHARPE, either directly as output from the program or after a modest amount of additional computation. The examples (in section 6) that illustrate the use of our method are made much clearer by adhering to the closed-form (SHARPE) framework; the reader can observe functions arising in the solution process explicitly.

For the remainder of the Introduction we examine the SHARPE methodology, closed-form solution, and hierarchical modeling in greater detail. Section 1 is a review of the basic concepts of stochastic process theory needed for the finite-state

semi-Markov processes that we deal with. Section 2 gives a form of the Chapman-Kolmogorov integral equations for a chain, in a manner that provides quantities necessary in the study of hierarchical decomposition. Section 3 presents the basic facts of the SHARPE solution method applied to certain types of chain. This is not to be interpreted as a literal description of code (the author is not a developer of SHARPE), but as background useful in judging what the program's capabilities are, and what enhancements might be desirable.

Section 4 presents an introductory example which shows how simple models can be built up into larger ones, and how repair complicates the decomposition issue. The problem is stated, of how to compute probabilities in a "combined model" (which may not even be semi-Markov). An integral equation is given whose solution answers this question. Section 5 is a digression on constructing certain exponomial distributions, and section 6 provides two examples that illustrate the power of our method.

Exponomial Distributions

To construct these distributions in an algebraic manner, we take as base field $\mathbb{R}$, the real numbers (an idealization of computer floating-point numbers). The set of functions $\{e^{\alpha t}, \sin \alpha t, \cos \alpha t, t\}$ where $\alpha \in \mathbb{R}$, generate a ring of functions, which is extended by linearity over $\mathbb{R}$ to form an algebra Exp. The subset of this algebra consisting of distribution is called $Dexp$, the *exponomial distributions*.

A function $F \in Dexp$ has the properties that

(1) $F(0) = 0$,
(2) $\lim_{t\to\infty} F(t) = 1$,
(3) $F(t)$ is non-decreasing for $t \geq 0$.

Condition 2) can be relaxed to $\lim_{t\to\infty} F(t) le 1$ for certain applications. In that case we say that F is defective, or incomplete, and has $Mass = 1 - \lim_{t\to\infty} F(t)$ at ∞. It is true that, given a finite state Markov process (chain) M, its unreliability function is a complete exponomial distribution provided that a certain technical condition holds. The condition is that every absorbing state is a failure state, and that whenever a state is exited, there is a non-zero probability that it will never be visited again. Hence the "operational" states are transient: call this the "transient state condition". If we weaken this condition to say merely that every failure state is an absorbing state, then the unreliability function is a defective exponomial distribution. This defines a larger class of functions, which we call $Dexp+$.

The SHARPE program calculates such an unreliability function in closed form for any Markov process satisfying the "transient state" criterion. In addition, SHARPE can find the unreliability distribution of certain semi-Markov processes. The process must be *acyclic* and the *transition functions* should be in $Dexp$. In a semi-Markov process, transition rates from one state (the present state) to another (the receiving state) are not constant, but depend upon the time elapsed since the system entered the "present state". We will later show (in section 3) how to use SHARPE to obtain more information about the chain (Markov or semi-Markov process) than is simply provided by the output of the program. For example, given a Markov process with cycles, SHARPE does not furnish the probability function (not a distribution) of a transient state. We show how to modify the chain, and then apply the C-K equations to find this. It is necessary only to solve one convolutional integral equation, not a system of equations, and to perform a few simple manipulations

using output generated by SHARPE. The integral equation is generally easy to solve using the Laplace transform. This approach was first used by Lotka in the theory of industrial replacement, where similar renewal-type integral equations also arise. See [Lotka]. The rigorous foundations of this solution method were brought together in [Feller2], using analytic techniques developed in [Churchill].

Hierarchical Modeling

As already indicated, hierarchical modeling methods are built into SHARPE. For instance, one way to construct a semi-Markov process is by means of state transition functions (distributions) given by the failure distribution of a (constant-rate) Markov chain. That is, in order to know explicitly a transition function of the semi-Markov chain, the "low-level" constant-rate chain must be solved. One process is in a sense embedded in the other. For examples of this type see [Sahner & Trivedi2]. A "full model" could be constructed by expanding the states in the higher-level semi-Markov chain into several states of a Markov chain. In many cases this capability of solving the higher- and lower-level models separately results in less computation and greater numerical robustness.

Another common class of decomposable systems consists of the Cartesian products. Given M_1, M_2 Markov chains, then $N = M_1 \times M_2$ is Markov. Selecting a state $\Lambda = (A, B) \in N$, $A \in M_1$, $B \in M_2$, one may modify N by removing all exit transitions from Λ to form $\tilde{N}$. Thus Λ is now an absorbing state. Define P_Λ to be the distribution function (possibly defective) corresponding to Λ. This function could be interpreted as "the probability that by time T we have simultaneously been in state A of M_1 and state B of M_2". We give several examples of how this scenario can arise in practice. The method we present allows one to find P_Λ without having to solve the large chain $\tilde{N}$. It is only necessary to obtain certain information about the smaller chains M_1 and M_2, which can be found for example by using SHARPE. Then another integral equation, (4.7), must be solved, leading to the desired probability function P_Λ. This integral equation is similar to the Chapman-Kolmogorov *forward* equation (see [Feller1] p. 458), and may be solved by the Lotka-Feller method.

This decomposition approach has several advantages. Work involved in solving an arbitrary chain with n states increases as the cube (n^3). In the Markov case, this is essentially the work involved in finding the eigenvalues of an $n \times n$ matrix. Hence, where decomposition is possible, much less work is needed to find the desired probabilities. Furthermore, the Cartesian product of even a Markov chain M_1 with S_1, a semi-Markov chain, need not be semi-Markov. In general, it is a stochastic process where transition rates depend on the time elapsed since entry into the state previous to the present state. Provided that one does not wish to deal with the theory of such general processes, the decomposition method is essentially the only viable approach.

SHARPE is a powerful tool for obtaining exponomial solutions of reliability models. The techniques suggested in this paper greatly expand the computational horizons of SHARPE in directions consistent with the hierarchical modeling philosophy, and thus have more than theoretical value. But the techniques may also be used without recourse to SHARPE, and using them in the context of numerical methods instead of closed-form solutions is a promising possibility as well.

1. Stochastic processes and distributions

We give a brief review of stochastic processes, with emphasis on the ones that are of greatest interest to us, namely *Markov* and *semi-Markov* processes (or chains). Although the term *chain* is sometimes used for discrete-time systems, we use the term to denote a continuous-time process that is either Markov or semi-Markov. The probabilistic definition of a Markov chain proceeds as follows. One starts with a space of outcomes, or *sample space* S. A continuous-time *stochastic process* is, for each $t > 0$, a function $X(t)$. The domain of the function is the sample space, and each image $X(t)\sigma$, where $\sigma \in S$, is an integer j from $1, \ldots, N$, identifying a state. So one may say that at time t, the process or outcome σ results in state j. It is most common to classify outcomes into events, and to assign probabilities to the various events. For a finite sample space where all outcomes are equally probable, of course

$$P(E) = \frac{\#\ of\ \sigma\ in\ E}{\#\ of\ \sigma\ in\ S},$$

and the expression (which is a slight abuse of notation) $P[X(t) = j]$ is the probability of the *event* consisting of all outcomes σ such that $X(t)\sigma = j$. The quantity $P[X(t) = j]$, called the *state probability* for state j. One can define a **finite Markov chain** to be a finite-state stochastic process X such that, for any set of times $t_0 < t_1 < \cdots < t_n < t$, the conditional probability that $X(t) = x$ given that $X(t_n) = x_n$, $X(t_{n-1}) = x_{n-1}, \ldots, X(t_0) = x_0$, where $x, x_0, \ldots, x_n$ are certain states, is equal to the conditional probability that $X(t) = x$ given that $X(t_n) = x_n$. This is the characteristic *memoryless property*. This finite Markov chain has the additional property of *time-homogeneity* if the quantity

$$p_{ij}(t) = P[X(u + t) = j | X(u) = i]$$

depends only on t and not on u, for all $u > 0$, i, j such that $i \neq j$. As usual, $P[E|F]$ means the probability of E given F, or $\frac{P[E \cap F]}{P[F]}$.

Now if one defines

$$\lambda_{ij} = \frac{d}{dt} p_{ij}(t)|_{t=0},$$

this quantity may be interpreted as follows. The increase over a short interval dt, of the probability of being in state j, *due to* the original probability of being in i, is equal to $P[X(t + dt) = j$ and $X(t) = i]$. From the definition of λ_{ij}, this equals $\lambda_{ij} P[X(t) = i] \cdot dt$.

By a suitable modification of what we have just done, we may obtain the definition and some properties of a semi-Markov chain. We use the idea of conditional probability density function, for which see [Trivedi]. Say that the process X *enters* state x_j at time t_j if $X(t_j) = x_j$, and if there is an $\epsilon > 0$ such that $X(t_j - \delta) \neq x_j$, for all $0 < \delta \leq \epsilon$. Also, if $X(0) = x_0$, the process X entered x_0 at $t = 0$ (x_0 was the *initial state*). Given a set of times $t_0 < t_1 < \cdots < t_n < t$, consider the conditional density function of X entering x at t given that X entered x_n at t_n, X entered x_{n-1} at $t_{n-1}, \ldots, X$ entered x_0 at t_0. If this is always equal to the density of X entering x at t given that X entered x_n at t_n, the process is said to be semi-Markov. We have for each pair i, j of states a density function

$$G_{ij}(t, u) = \text{density of } X \text{ entering } j \text{ at } u + t \tag{1.1}$$

conditional on X entering i at u.

The time-homogeneous case occurs when G_{ij} is independent of $u > 0$ for all $i, j,\ t > 0$. We shall henceforth refer to a time-homogeneous Markov or semi-Markov process as a *chain*. Such a chain, since Markov implies semi-Markov, is characterized by the functions

(1.2) $\qquad G_{ij}(t) =$ density of X entering j at t given that

X entered i at time 0.

It has been shown (see [Ross], p. 89) that certain other sets of functions serve to characterize a chain. Consider the (possibly defective) *distribution* $F_{ij} = P[X$ enters j at some time τ, $0 < \tau \leq t$, and X does not enter any state at any time κ, $0 < \kappa < \tau | X$ entered i at 0]. In words, F_{ij} is the probability, conditional on entering i at 0, of ending the sojourn in i by a *jump to* j before time t. In his 1964 study [Feller3] of the C-K equations, Feller makes use of F_{ij}. Another class of functions which determine a chain is referred to as the *transition distributions* C_{ij}. Described in words, $C_{ij}(t)$ is the probability that X will jump to j (first entry) by time t, given that X entered i at 0 and assuming that C_{ij} is the only transition out of state i. Thus C_{ij} is a distribution valid in the absence of competing transitions. The distributions $\{C_{ij}\}$, over all j, are assumed to correspond to the *independent events* of jumping from i to the various states j. The functions $\{C_{ij}\}$ are what is supplied to SHARPE when it is desired to "solve" a chain (determine the time-dependent probability functions of its states).

A relation between F_{ij} and C_{ij} will be given subsequently. For instance, when the chain is Markov, we have

(1.3) $\qquad C_{ij}(t) = 1 - e^{-\lambda_{ij} t}$, and

$$F_{ij}(t) = \frac{\lambda_{ij}}{\sum_k \lambda_{ik}} \cdot (1 - e^{-\sum_k \lambda_{ik} t}).$$

The C_{ij} functions can be given in various ways. For certain purposes, it is not necessary to define them completely but rather it is enough to give their mean and variance as distributions. This approach is used by the SURE [Butler & White] package to find upper and lower bounds on the reliability of a system whose reconfiguration times are not exponentially distributed. The SHARPE package, on the other hand, expects to be provided with C_{ij} as a function in the class *Dexp*. In other words,

(1.4) $$C_{ij}(t) = \sum_{r=1}^{m} a_r t^{k_r} e^{b_r t},$$

where k_r is a non-negative integer and a_r and b_r are real or complex numbers. C_{ij} should be a complete distribution function, in particular real-valued; from this it is not hard to show that the terms with non-real coefficients a_r, can be matched in pairs with indices r, r', such that $k_r = k_{r'}$, $a_r = \overline{a_{r'}}$, $b_r = \overline{b_{r'}}$, where the bar denotes complex conjugation. This property will be referred to as the "conjugacy

condition". A typical expression would be

$$1 - e^{-t} + \left(\frac{ia}{2}\right)\left[e^{-(1-i)t} - e^{-(1+i)t}\right] \tag{1.5}$$

or $1 - e^{-t} - ae^{-t}\sin t$.

Having made the requisite definitions, we introduce standard terminology relating to the classification of chains in order to simplify later exposition.

Definition 1.6 A chain is *ergodic* if, given that it is in state j at time t, if k is another state, then there is a later time t_k such that $P[X(t_k) = k] > 0$.

Definition 1.7 A state k is *absorbing* if, given $X(t) = k$, then $P[X(t') = j] = 0$, for all $t' > t$ and $j \neq k$. Thus an absorbing state, once entered, can never be left. Similarly, a subset of the set of states could form an *absorbing subchain* if once entered, it is never left. Clearly, a chain with an absorbing subchain that is not the whole chain *cannot* be ergodic.

Definition 1.8 A state k is *transient* if there is a state $j \neq k$ such that given $X(t) = k$, then for some $t' > t$, $P[X(t') = j] > 0$, but given $X(t) = j$, then for all $t' \geq t$, $P[X(t') = k] = 0$.

Definition 1.9 A chain is *irreducible* if it has no absorbing subchain, other than itself.

The class of exponomial distribution functions is a natural one for the study of chains. In fact, in the Markov case, the function P_A as above is exponomial as can be seen from the differential theory of constant-rate chains. Let Q be the "infinitesimal generator" matrix, $q_{ij} = \lambda_{ij}$ for $i \neq j$, $q_{ii} = -\sum_{j \neq i} \lambda_{ij}$. Then if $\vec{P}(t)$ is a row vector of functions $[P_1, \ldots, P_i, \ldots, P_n]$, where the n states of the chain are numbered $1, \ldots, n$ and $P_i(t)$ is the probability of being in state i $P[X(t) = i]$ at time t, we have

$$\vec{P}'(t) = \vec{P}(t)Q, \quad \vec{P}(0) = \vec{P}_0, \tag{A1}$$

[Reibman Trivedi]. Here $\vec{P}_0$ is the vector of initial probabilities. Then

$$\vec{P}(t) = \vec{P}_0 e^{Qt}, \tag{A2}$$

where we use the matrix exponential $e^{Qt} = \sum_{i=0}^{\infty} \frac{(Qt)^i}{i!}$. Putting Q into Jordan normal form $Q = SJS^{-1}$, it follows from [Moler & Van Loan], p. 381] that

$$\vec{P}(t) = \vec{P}_0 S e^{Jt} S^{-1}. \tag{A3}$$

If $J = \mathrm{diag}(J_1, \ldots, J_p)$, then

$$e^{Jt} = \mathrm{diag}(e^{J_1 t}, \ldots, e^{J_p t}). \tag{A4}$$

If

$$J_i = \begin{bmatrix} \lambda_i & 1 & 0 & 0 & 0 \\ 0 & \lambda_i & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & 0 \\ \vdots & \vdots & \cdots & \ddots & 1 \\ 0 & 0 & \cdots & \cdots & \lambda_i \end{bmatrix}$$

and $m_i \times m_i$ complex matrix, then If

$$e^{J_i t} = \begin{bmatrix} e^{\lambda_i t} & te^{\lambda_i t} & \cdots & \cdots & \frac{t^{m_i-1}e^{\lambda_i t}}{(m_i-1)!} \\ 0 & e^{\lambda_i t} & \cdots & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \cdots & \ddots & te^{\lambda_i t} \\ 0 & 0 & \cdots & \cdots & e^{\lambda_i t} \end{bmatrix} \tag{A5}$$

From formulas A1-A5 it follows that any $P_i(t)$ can be written

$$P_i(t) = \sum_{j=1}^{m} \sum_{k=0}^{p_i-1} a_{ijk} t^k e^{-\lambda_j t}, \tag{1.10}$$

where m is the number of distinct eigenvalues of Q, λ_j is the j-th distinct eigenvalue of Q and p_j is the multiplicity of the factor $(x - \lambda_j)$ in the minimum polynomial of Q. Since $P_i(t)$ must be a real function, the conjugacy condition must hold, and we have an exponomial function. If i is an absorbing state of a Markov chain, P_i must be a distribution function, but may be defective. If i is the only absorbing state, then P_i is a complete distribution, under the assumption made above that all states are either absorbing or transient.

2. Chapman-Kolmogorov Equations

We present a form of the Chapman-Kolmogorov equations for a semi-Markov process which will be convenient for our purposes. These are a form of the *backwards* equations in that they give probabilities (and densities) by summing over all epochs (times of a jump), and all results of a jump out of a given state, for the *first jump* from that state. Similar equations are stated, with proof, in [Ross], p. 93. We deal with state probability functions, and their "densities", which integrate to give the probability function. Thus a density does not have to be the derivative of a *distribution*.

We recall some of our semi-Markov terminology from section 1. In this section, i, j, k are states; then $C_{ik}(T)$ = probability that a jump from i to k would be made by time T, given that i was entered at time 0, in the absence of competing transitions. These are the transition distributions, and they are assumed to be independent and competing. (They are distributions of the independent events resulting in jumps to the different states.) This is the same as the definition from section 1 provided that C gives a complete distribution. Recall that the unconditional transition function $F_{ik}(T)$ is the possibly defective distribution of a jump from i to k by time T, given

that i was entered at time 0. The two distributions are related by

$$F_{ik}(T) = \int_0^T C'_{ik}(t) \prod_{j \neq k} [1 - C_{ij}(t)] dt. \tag{2.1}$$

In words, the probability of leaving i for k by T is the integral of the density of jumping from i to k at t, times the probability of not having jumped to any other state by t.

Next let $E_{ik}(t)$ = density of (first) entry time from i to k, given that i was entered at 0. This is the density corresponding to a possibly defective probability distribution. We have

$$E_{ik}(T) = dF_{ik} + \sum_{j \neq k} \int_0^T dF_{ij}(\tau) \cdot E_{jk}(T - \tau), \quad i \neq k, \tag{2.2}$$

$$E_{kk}(T) = \sum_{j \neq k} \int_0^T dF_{kj}(\tau) \cdot E_{jk}(T - \tau).$$

In words, for the first equation, the density of *first* arrival in k is the density of *jumping* to k plus the density which results from jumping to a third state at a time $\tau \leq T$, followed by a subsequent first arrival in k at T.

The density $G_{ik}(t)$ defined above in (1.2) will be used in forming state probability functions, and is necessary to compute probabilities of "combined states" in hierarchical, or Cartesian product, models. We recall that it is the density of entering k at t, given that you entered i at 0. This differs from E_{ik} in that state k may previously have been visited (after leaving state i). Then we have

$$G_{kk}(T) = E_{kk}(T) + \int_0^T E_{kk}(\tau) G_{kk}(T - \tau) d\tau \tag{2.3}$$

$$G_{ik}(T) = E_{ik}(T) + \int_0^T E_{ik}(\tau) G_{kk}(T - \tau) d\tau$$

The first expression is a Volterra integral equation of the second kind, for which [Burton] is a clear introductory source. It is the same type of equation discussed in Feller's article on renewal theory [Feller2] and which was previously treated as part of the theory of industrial replacement in [Lotka]. The equation is also pivotal in later studies of population and economic growth, as more recent references from chapter 2 of [Kohlas] indicate.

The verbal description of the equation and formula above are now given. For the equation, "given that you start in k, the density of entering k at T is the density of a first entry, plus the density of having entered k at a previous time τ and subsequently entering at T, summed over all τ". For the formula, "starting in j, the density of arriving in k is the sum of the density of first arrival, plus that of a previous first arrival followed by a subsequent arrival from k to k".

These quantities can be used to express the state probabilities. That is, $P_{ik}(T)$ is the probability of being in k at T, given that you entered i at 0. We let $S_k = \sum_{j \neq k} F_{kj}$

be the holding time distribution in state k. Then

$$P_{ik}(T) = \int_0^T G_{ik}(\tau)[1 - S_k(T-\tau)]d\tau, \tag{2.4}$$

$$P_{kk}(T) = 1 - S_k(T) + \int_0^T G_{kk}(\tau)[1 - S_k(T-\tau)]d\tau.$$

For the first formula, a description in words reads: "the probability of being in k equals the density of arriving in k, and subsequently not leaving k, integrated up to the present time".

An alternative formulation of state probabilities, using only the E_{ij} functions can be derived from the above equations. In fact,

$$P_{ik}(T) = \int_0^T E_{ik}(\tau)P_{kk}(T-\tau)d\tau, \tag{2.5}$$

$$P_{kk}(T) = 1 - S_k(T) + \int_0^T E_{kk}(\tau)P_{kk}(T-\tau)d\tau.$$

Again, the advantages to the approach indicated by the above formulas can be summarized:

(1) The method encompasses both Markov and semi-Markov chains,
(2) E_{ik}, $i \neq k$ can be found by SHARPE,
(3) after which only one integral equation need be solved to determine a probability function,
(4) then Feller's method of solving the renewal equation can be applied;
(5) the approach is well adapted for hierarchical modeling as will be seen.

3. The SHARPE Solution Method

Acyclic chains (Markov and semi-Markov)

The method adopted by SHARPE in this case is equivalent to an analysis of paths from "initial states" to absorbing states. An initial state can be defined as one that has a non-zero probability at time 0. That is, if i is the state, then the vector $\vec{P}_0$ has a positive i-th component. We discuss this by examining each path separately, as in a "depth-first search", whereas SHARPE is actually programmed to compute probabilities at states as they are successively reached in a "breadth-first" search. The difference is one of form.

For the system to traverse a particular path in the (acyclic, directed) graph representing the semi-Markov process is an *event*, which is disjoint from the other events corresponding to the other paths. Therefore, to get the distribution of an absorbing state, the traversal distributions of all paths leading from some initial state must be added, weighted by their probabilities of occurrence. We must find a traversal distribution from a given path, and a probability. Suppose the initial state is i_0, the final state is i_m. The path of concern σ can be written $i_0, i_1, \ldots, i_m$. We define recursively

$$B^m(T) = 1, \quad B^j(T) = \int_0^T F'_{i_j,i_{j+1}}(t)B^{j+1}(T-t)dt, \tag{3.1}$$

where $j = 0, \ldots, m-1$. Then define $D_\sigma = B^0(T)$. The derivative and the convolutions are performed "symbolically" by SHARPE, within the class of functions *Exp*.

Let $\rho_\sigma = \rho_{i_0,i_1} \cdot \rho_{i_1,i_2} \cdots \rho_{i_{m-1},i_m} \cdot P_{i_0}(0)$. Here,

$$\rho_{i_j,i_{j+1}} = \int_0^\infty F'_{i_j,i_{j+1}}(t)dt, \quad j = 0, \ldots, m-1.$$

Then $P_k(T) = \sum_\sigma \rho_\sigma D_\sigma(T)$, where σ runs over all paths ending in k.

Cyclic chains (Markov)

In the (cyclic) Markov case, SHARPE uses both matrix analysis and estimation in the transform domain to find the distribution of an absorbing state. There is no reason why this method could not be used to give probability functions at transient states, but at present, SHARPE does not do this. Such information may be useful, however, as the example of a *phased mission* points up. Here the "mission" proceeds in two phases, the second commencing at time T_1. The models for the two phases are the same, but certain failure and reconfiguration distributions have changed due to a maintenance action ([Baker], *personal communication*). The initial state probabilities for the model of the second phase are given by the state probabilities of the first phase at time T_1. We indicate how to use SHARPE to find this information, however.

Recalling the infinitesimal generator matrix Q of section 1, it is clear from (1.10) that its eigenvalues and their multiplicities are of great importance in finding the probability functions. Any real matrix has a *Schur decomposition*

$$Q = UHU^T, \tag{3.2}$$

where U^T denotes the transpose of U, U is orthogonal ($UU^T = I$). The $n \times n$ matrix H is to have a nearly upper triangular form. That is, it is *block upper triangular*, with diagonal blocks either of size 1×1 or 2×2. In particular, H is an *upper Hessenberg* matrix: it is upper triangular except for possible non-zero entries on the diagonal $i = j + 1$ (just below the main diagonal). Then the eigenvalues of H, and hence of Q are the 1×1 real scalars and the complex-conjugate pairs arising from the 2×2 blocks. In order to take Q to this form, one may first find

$$G = L_k L_{k-1} \cdots L_0 \cdot Q \cdot L_0 \cdots L_k,$$

where G is in upper Hessenberg form, and L_i is a "Householder matrix". A Householder matrix represents a reflection through an $(n-1)$-dimensional hyperplane orthogonal to a certain vector $\vec{v}_i$. The details of this algorithm can be found in [Golub & Van Loan] p. 222. It is due to Wilkinson who exposited it in his book [Wilkinson]. Now G has a *Q-R decomposition* $G = WR$, where W is orthogonal (essentially a product of rotations) and R is upper triangular. Another algorithm implicitly finds $G' = RW$. Using G' as the new Hessenberg matrix G, we repeat this process until the real Schur form is attained. Then the eigenvalues (which have not changed through any of these transformations) may be read off.

Next, SHARPE must determine the coefficients a_{ijk} of formula (1.10). By transforming the differential equation in formula A1 of section 1, one obtains

$$s\overline{P}(s) - P_0 = \overline{P}(s)Q,$$

or $\overline{P}(s \cdot I - Q) = P_0$. Each $\overline{P}_i(s)$ is of the form

$$\sum_{j=1}^{m} \sum_{k=0}^{p_i - 1} \frac{\beta_{ijk}}{(s + \lambda_j)^k};$$

determining the β_{ijk} is equivalent to finding a_{ijk}. But for a particular choice of s, say ζ_1, we get $\overline{P}(\zeta_1)T = P_0$, $T = (\zeta_1 I - Q)$. Thus we have n equations for the n^2 unknowns $\{\beta_{ijk}\}$. Similarly, setting $s = \zeta_2, \ldots, \zeta_n$, for suitably chosen values, will give enough equations to determine the coefficients we seek.

We now indicate how to use the SHARPE approach to determine transient state probabilities. This will work for any Markov chain; if general (semi-Markov) transitions are allowed, the technique is only good for states (if any) that satisfy the following. "If all transitions *out* of the state of interest are removed, that which remains is a chain that is

(1) pure Markov, possibly with cycles, or
(2) acyclic semi-Markov."

Call this condition Condition Q. It is rather remarkable that SHARPE, with some additional calculation, can treat certain semi-Markov chains with cycles. The computational techniques involved are amply illustrated by the examples at the end of the paper. Now the method is described in general terms.

Given a (non-absorbing) state r, we are interested in $P_r(T)$ as a function. If there is a single initial state j, $P_j(0) = 1$, this is the same as $P_{jr}(T)$. But using SHARPE, for any state $i \neq r$, one may find $E_{ir}(T)$. This is done by describing the chain to SHARPE, giving the transition rates and distributions as usual, but omitting any transitions out of r. This makes r into an absorbing state r'. We assign in the input to SHARPE, $P_i(0) = 1$, and all other initial state probabilities zero. Given that Condition Q holds, SHARPE can find the cumulative distribution function $H_{r'}(t)$ of arrival into r', as well as the overall probability $\rho_{r'}$ of reaching r'. Consider

$$L_{ir'}(T) = H_{r'}(T) \cdot \rho_{r'}.$$

This is the unconditional distribution of entering r'. The derivative of $L_{ir'}$ with respect to time is just $E_{ir}(T)$, since arrival and first arrival are identical for an absorbing state. Thus SHARPE has found the functions E_{ir}, $i \neq r$. These are then used by means of the second part of (2.2) to find $E_{rr}(T)$. The second part of (2.5) is an integral equation for the unknown function P_{rr}. Once this has been solved, we need only perform the convolution integration of (2.5), first equation, to obtain P_{jr} which was the desired state probability function.

4. Decomposition Methods

Given a complex failure-repair-reconfiguration system, it is tempting to decompose it into independent subsystems, or ones that are nearly independent. Independence allows one to compute the probability of being in a given state (for each

subsystem) by using the product formula. Since the computational cost of analyzing a system model increases geometrically with size, significant savings can be had if the system is decomposable in this manner. As an example, consider two triplex systems attached to a voter as shown in Figure 4-1. are both required to be functioning for system viability. In each subsystem, S_a, S_b, the failure of two components causes a triplex failure. If the component failure probabilities are p_a and p_b, we have

$$p_S = p_{S_a} + p_{S_b} - p_{S_a} \cdot p_{S_b}, \tag{4.1}$$

where

$$p_{S_a} = 3 \cdot p_a^2(1 - p_a) + p_a^3,$$
$$p_{S_b} = 3 \cdot p_b^2(1 - p_b) + p_b^3.$$

Thus we see that in forming system failure probability, we certainly do not need to consider separately all failure modes, such as "one unit in S_a has failed, together with 2 units in S_b".

Independent Voted Triplexes

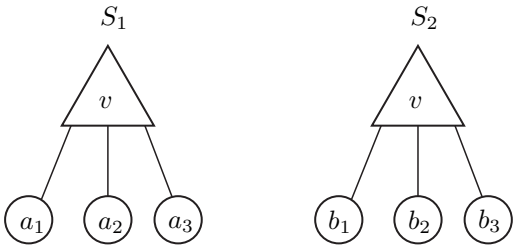


Figure 4-1

Dependent Processor-Node System

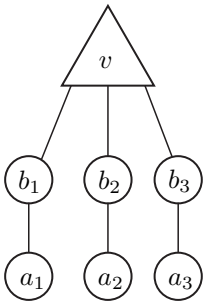


Figure 4-2

On the other hand, in Figure 4-2. the same failure conditions apply for S_a and S_b, but their "failures" are not independent events. We say that unit B_1 is "isolated" when a_1 has failed (the voter has no access to it). To be isolated is as

bad as failed. Thus when a_1 and b_2 are failed, the system has failed, since the voter can see neither b_1 nor b_2. The "failure conditions" need to be explicitly analyzed. If we take the simplest Markov chain representation of S_a and S_b, we get Figure 4-3. At a given time t we again have

$$P_{S_a}(t) = 3 \cdot P_a^2(t) - 2 \cdot P_a^3(t) \tag{4.3}$$

for the failure probability. But it is not clear how to obtain $P_S(t)$ for the combined system. Figure 4-4a gives an "equivalent" Markov chain; its failure probability function is the same as for Figure 4-3. The corresponding chain for subsystem S_b is shown in Figure 4-4b. In the "combined" model (not shown), certainly when one of the subsystems is in a failed state, the system is failed. Thus $(F_a, *)$ and $(*, F_b)$ give system failure, where $*$ is any non-failed state of the appropriate subsystem. But also for example $(011, ioi)$ is a failed state. Due to independence of unit failures, this state's probability is $P_{011} \cdot P_{ioi} = P_a(t) \cdot P_b(t)$. Examination gives 6 of these states so we finally get

$$P_S(t) = P_{S_a}(t) \cdot (1 - P_{S_b}(t)) + (1 - P_{S_a}(t)) \cdot P_{S_b} + 6 \cdot P_a(t) \cdot P_b(t). \tag{4.4}$$

Voted Triplex System

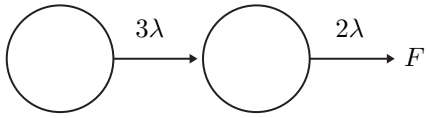
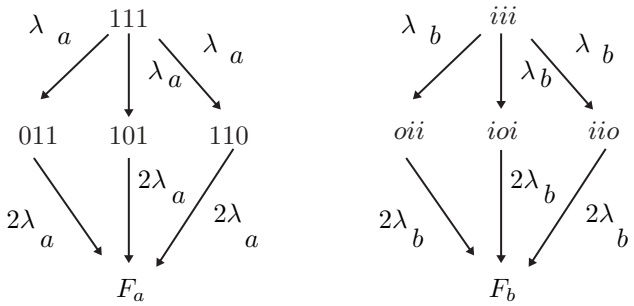


Figure 4-3

Elaborated Triplex Models



(a) (b)

Figure 4-4

Therefore, the combined model is the Cartesian product $S_a \times S_b$ with certain transitions modified. For example, the definition of Cartesian product implies that $(101, ioi)$ is a state with a transition of rate $2\lambda_a$ to (F_a, ioi) and a transition of rate $2\lambda_b$ to $(101, F_b)$. Next, all "combined states" satisfying the failure condition are made absorbing. Thus the transitions of rate $2\lambda_a$ from $(110, ioi)$ to (F_a, ioi) and

rate $2\lambda_b$ from $(110, ioi)$ to $(110, F_b)$ are deleted. This reliability problem, of coupled nodes and sensors, can be solved in two ways: firstly by forming and solving the combined Markov model in the manner we have just indicated, and secondly by solving each of the two models S_a and S_b, not only for their failure probability distributions, but also for their state probability functions. These functions are then combined in some way, similar to (4.4), to give the distribution of the entire system S.

The situation becomes more involved when the components admit of repair. The Markov model for system S_a is then shown in Figure 4-5, and the model for S_b is similar. At time T, if we are in a failed state of S_a or of S_b, the system S has certainly failed. If we are in a state such as 011 in S_a and ioi in S_b, the system is failed as well. But we may well be in an "up" state, such as 111 in S_a and iii in S_b and still have to consider that we are *failed*. This is because at *some previous time* $t < T$, we may have been in 011 and ioi simultaneously which would have brought down the system.

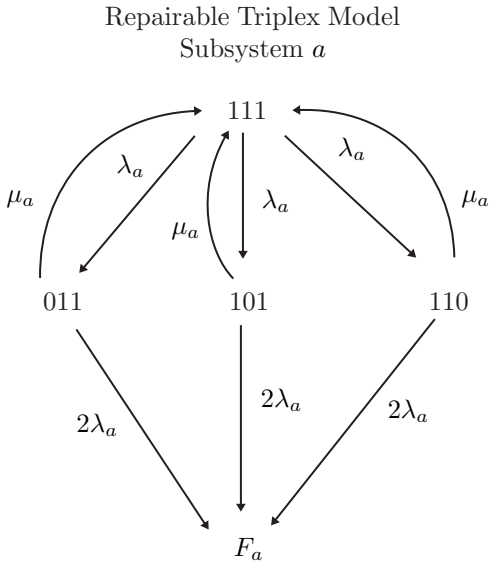


Figure 4-5

The combined model, called **m** accurately reflects this state of affairs: the state $(011, ioi)$ has been made absorbing, so the transitions (named after their numerical rate):

$$\begin{aligned} \mu_a &: (011, ioi) \to (111, ioi) \\ \mu_b &: (011, ioi) \to (111, ioi) \end{aligned} \tag{4.6}$$

do not exist.

The problem is how to compute the failure contribution of combined states such as $(011, ioi)$ without solving the combined model **m**. This is analogous to the non-repair situation, with the difference that we cannot simply use the expression $P_{011}(T) \cdot P_{ioi}(T)$ as we did there. The expression we seek could be expressed in

words as "the probability that at some time prior to T, the S_a state was 011 and the S_b state was ioi".

In the semi-Markov case it is not feasible to find these "combined state" probabilities by using a Cartesian product model. If M and N are semi-Markov chains, the Cartesian product $M \times N$ will generally not have the semi-Markov property. For example, Figure 4-6 depicts a two state Markov chain and a two state semi-Markov chain with hypoexponential distribution $C(t) = 1 - 2 \cdot e^{-t} + e^{-2t}$. The combined model (Cartesian product) is then shown with distributions indicated. Since $C(t)$ is not exponential and hence not memoryless, the density function of the transition $(b, x) \to (b, y)$ depends not only on the "local" time spent in state (b, x), but also on the entry time into (b, x), or the time spent in (a, x). This violates the semi-Markov property.

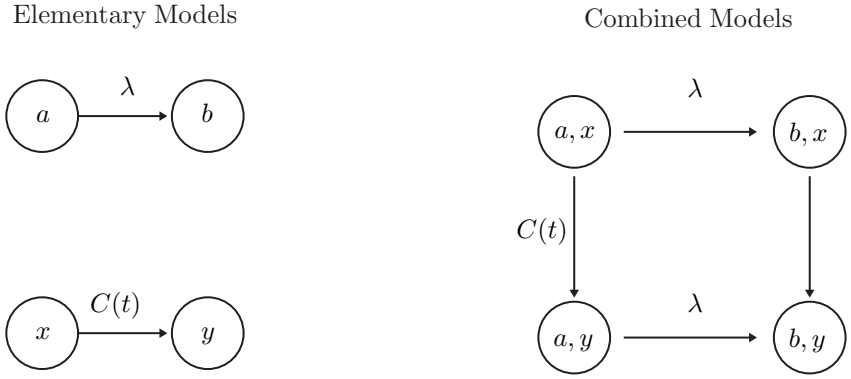


Figure 4-6

For this reason we do not work explicitly with the "combined model". But we still consider ordered pairs of states, and say, informally, "the system is in state (A, X), where A is a state of M, and X is a state of N". We a consider failure condition given by a pair (B, Y), $B \in M$, $Y \in N$. Let $Z_{AB,XY}(T) =$ the probability that M has entered state B while N was in state Y, or N entered state Y while M was in state B, at a time t, $0 < t < T$, given that M was in A at 0, and N was in X at 0. It should be helpful to look ahead to Figure 6-1 which gives a good illustration of this situation.

In case $A \neq B$ or $X \neq Y$, one can also interpret $Z_{AB,XY}(T)$ as follows: the probability, given that M started in A and N started in X, that M has been in B simultaneous with N being in Y. The quantities are determined by means of two fundamental equations. The first is:

(4.7)

$$Z_{BB,YY}(T) = \int_0^T (G_{BB}(\tau) \cdot P_{YY}(\tau) + P_{BB}(\tau) \cdot G_{YY}(\tau))[1 - Z_{BB,YY}(T - \tau)]d\tau.$$

In words, the right-hand expression is the integral over τ of the density of entering into the "state" (B, Y), and subsequently never arriving again (to avoid counting arrivals twice). This is similar to a Chapman-Kolmogorov *forward* equation in that we integrate over densities of the *last jump* into (B, Y). The quantities G_{BB}, P_{BB}, G_{YY}, P_{YY} are found as in section 2 from the separate models M and N. Then

(4.7) is an integral equation to be solved. Note that if we use a Laplace transform method, the expression $G_{BB}(\tau) \cdot P_{YY}(\tau)$ must be multiplied in the time domain, and then transformed, or else $\bar{G}_{BB}(s)$ and $\bar{P}_{YY}(s)$ *convolved* before proceeding further. This is illustrated in the subsequent examples. Given $Z_{BB,YY}$, one can find $Z_{AB,XY}$ by integrations:

(4.8)

$$Z_{AB,XY}(T) = \int_0^T (G_{AB}(\tau) \cdot P_{XY}(\tau) + P_{AB}(\tau) \cdot G_{XY}(\tau))[1 - Z_{BB,YY}(T - \tau)]d\tau.$$

The verbal interpretation of the right-hand expression is left to the reader.

5. Distributions from Mean and Variance

Modern fault-tolerant computers, as used in high-reliability applications such as aerospace and nuclear plant control, employ architectural features beyond simple majority voting of independent processors. Instead, faulty components may be switched off, and spares activated; the system is changed upon detection of a fault. A simple system with such dynamic reconfiguration is shown in Figure 5-1. This depicts the triplex degradable to a simplex mentioned in section 1. Practice has generally borne out the constant failure rate assumption for electronic components during their active life span. But the "reconfiguration distribution" $\omega(t)$ has been observed not to be exponential, as in [Finelli]. This transition includes the time necessary for the system to detect the presence of single fault, isolate the two components (one good and one bad), and remove them from service.

Reconfigurable Triplex

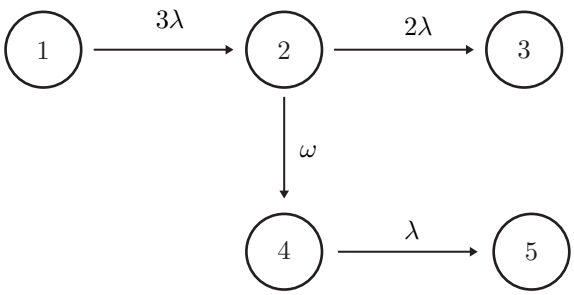


Figure 5-1

It has been shown in [Butler & White] that giving the mean M and variance V of $\omega(t)$ is sufficient to determine $P_S(T)$ to within a few percent, assuming that M is much smaller than the reciprocal of the largest failure rate in the system S. Here, T is the mission time. That is, the system is assumed to fail slowly and reconfigure quickly.

To check such reliability results, obtained by the SURE program package, one might use SHARPE on the same example. To do so would necessitate presenting $\omega(t)$ in exponomial form. Our goal in the present section is simply to give a way of determining $\omega(t)$ explicitly, knowing that it is a distribution with mean M and

variance V. We utilize the method of Cox from his classic paper [Cox]. Three cases exhaust the possibilities.

Case 1. Suppose $M^2 = V$. Then take $\omega(t) = 1 - e^{-\lambda t}$, where $\lambda = 1/M$.

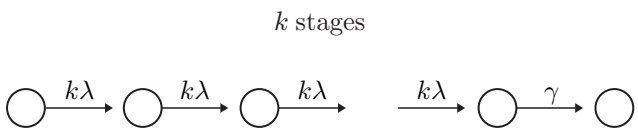


Figure 5-2

Case 2. If $M^2 > V$, let $k = [M^2/V]$ be the greatest integer less than M^2/V. Consider the linear chain in Figure 5-2, consisting of k stages with rate $k\lambda$ and a final stage of rate γ. Since the random variable "time to failure" is the sum of the independent transition times of the stages, the mean and variance are additive (respect the summation). See [Trivedi], p. 192. Thus

$$M = k\left(\frac{1}{k\lambda}\right) + \frac{1}{\gamma} = \frac{1}{\lambda} + \frac{1}{\gamma},$$

$$V = k \cdot \frac{1}{k^2\lambda^2} + \frac{1}{\lambda^2} = \frac{1}{k\lambda^2} + \frac{1}{\gamma^2}. \tag{5.1}$$

From this one obtains the formula

$$\frac{1}{\lambda} = \frac{M - \sqrt{M^2 - \left(1 + \frac{1}{k}\right)(M^2 - V)}}{\left(1 + \frac{1}{k}\right)}. \tag{5.2}$$

The practical way to get $\omega(t)$ in closed form is to find k, λ, γ and enter a SHARPE file for the linear chain. SHARPE will then find the desired distribution $\omega(t)$. Figure 5-3 shows the input and output formats. In a hierarchical fashion, SHARPE allows the cumulative distribution function of this chain to be used in a "higher" system, eliminating the need ever to write the exponomial form of $\omega(t)$ explicitly.

A complete explanation of SHARPE input and output formats should be found in [Sahner & Trivedi1]. Certain symbolic variable names such as "lambda" are *bound* to a numerical value. the system is described by *type* (markov) and given a name (linear). The states and transitions, with rates, are given in the following lines; after an "end", state 0 is assigned initial probability 1. Then the cumulative distribution function (cdf) is *requested* of SHARPE. The cdf then appears in the output, using mantissa and exponent notation to describe floating point numbers, and "exp" to denote the exponential function. Hence the meaning of the first line of the output is $8.5749 \times 10^{-2} t^3 e^{1.6668t}$. Finally the mean and variance of the cdf are given.

Case 3. If $M^2 < V$, we take $\omega(t)$ to be *hyperexponential* with distribution

$$\omega(t) = 1 - pe^{-\mu t} - qe^{-\lambda t}, \tag{5.3}$$

```
bind
lambda .4167
gamma .16667
end
markov linear
0 1 4*lambda
1 2 4*lambda
2 3 4*lambda
3 4 4*lambda
4 5 gamma
end
0 1.
end
cdf(linear)
end
CDF for system linear:

 8.5749e-02 t( 3) exp(-1.6668e+00 t)
+ 3.2582e-01 t( 2) exp(-1.6668e+00 t)
+ 6.1957e-01 t( 1) exp(-1.6668e+00 t)
+ 1.0000e+00 t( 0) exp( 0.0000e+00 t)
- 1.5241e+00 t( 0) exp(-1.6667e-01 t)
+ 5.2412e-01 t( 0) exp(-1.6668e+00 t)

mean: 8.3997e+00
variance: 3.7438e+01
```

Figure 5-3

where $p + q = 1$. According to [Trivedi], p. 212,

$$M = \frac{p}{\mu} + \frac{q}{\lambda}$$

$$V = \frac{2p}{\mu^2} + \frac{2q}{\lambda^2} - M^2. \tag{5.4}$$

An effective iterative procedure to find p, q, μ and λ is to set

$$\begin{aligned} &p = q = 5, \\ &\lambda = 1.1/M, \\ &\mu = 1/\left(2M - \frac{1}{\lambda}\right). \end{aligned}$$

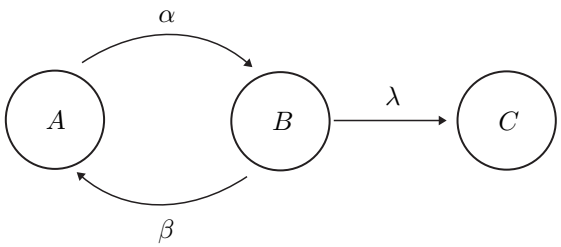
Then let (Step 1) $X = (M^2 + V)/2 - q/\lambda^2$. This should approximate p/mu^2. Then set a new μ value (Step 2) equal to $(M - q/\lambda)/X$. Multiplying out the following

expression shows (Step 4) that $\mu\lambda(M - 1/\lambda)/(\lambda - \mu)$ should equal p, so we take this value as our new p. Finally, (Step 5), take $q = 1 - p$, and begin again at (Step 1), repeating until the computed mean M and variance V are as close to the given values as needed. This method is used in the next section to construct a distribution.

6. Examples

As our first example we consider the two models I and II depicted in Figure 6-1. Model I is a transient-fault detection mechanism. In state A the mechanism is functioning normally; in state B transient faults are incorrectly diagnosed as being permanent. See [Lala] p. 20. In state C, a rare kind of error causes spurious signals to be sent external parts of the system, causing an overall crash.

Model I (Transient Error Detection)



Model II (Error Arrival and Recovery)

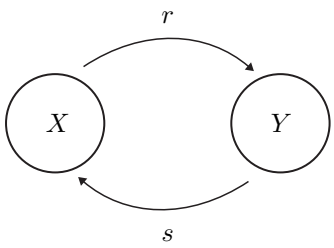
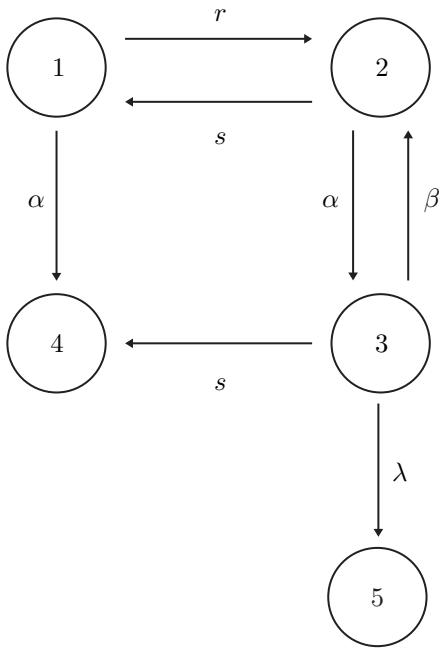


Figure 6-1

Model II represents the arrival of, and recovery from, transient faults over the entire system. State X represents the active presence of a fault, and Y the disappearance (absence), of faults. A similar model could be used to depict the error-producing and benign phases of a single "intermittent" fault. Several other reliability estimation packages besides SHARPE provide a capability for modeling the arrival and detection of permanent, transient, and intermittent faults to the system. See [Trivedi et al] and [Bavuso & Peterson].

We wish to consider the system as being *up* when the two "subsystems" I and II are in states (A, X), (A, Y), or (B, Y) respectively. Whenever model I is in state C, the system is down, but also whenever I is in state B at the same time as model II is in state X, we must consider the system to have crashed, since a transient fault is present but is incorrectly diagnosed (as permanent).

Combined Model



State Correspondence: $1 \sim A, X \quad 2 \sim A, Y \quad 3 \sim B, Y \quad 4 \sim B, X \quad 5 \sim C, Y$

Figure 6-2

The "combined model", with states $\{1, \ldots, 5\}$ is shown in Figure 6-2. The correspondence between the states of the combined model and the Cartesian product $ItimesII$ is indicated. The system failure states are absorbing. Note that there is no state corresponding to (C, X): it is superfluous. Thus, to find the failure probability at time T, one may take

$$P_4(T) + P_5(T). \tag{6.1}$$

We are most interested in $P_4(T)$, the probability of failure due to being in "state" (B, X) at some time $t \leq T$. For simplicity, suppose that the system starts life with model I in state A and model II in state X. In the notation of section 4, we see that

$$P_4(T) = Z_{AB,XX}(T). \tag{6.2}$$

The total failure probability can also be obtained from the solution of the individual models I and II. The remaining part to be considered is for state C to be entered while model II is in state Y. Since C is absorbing, we know that the density of entering C in model I is $G_{AC} = E_{AC}$, and we obtain

$$\int_0^T E_{AC}(\tau) \cdot P_Y(\tau) d\tau. \tag{6.3}$$

Then adding expressions (6.2) and (6.3) gives the total failure probability, and should be equal to the distribution obtained from considering the absorbing states 3 and 5 of the combined model.

In finding $Z_{AB,XX}(T)$ as a closed-form exponomial function, we will need to know G_{BB}, P_{BB}, G_{AB}, P_{AB}, G_{XX}, and P_{XX}, as indicated by equations (4.7) and (4.8). We set coefficients in system I as

$$\begin{aligned}\alpha &= .3,\\ \beta &= .5,\\ \lambda &= .1.\end{aligned}$$

We have $dF_{BA}(t) = .5e^{-.6t}$, and since there is only one transition into B, it also follows that

$$E_{AB} = dF_{AB} = .3e^{-.3t}.$$

Thus by (2.2),

$$E_{BB}(T) = \int_0^T dF_{BA}(\tau) \cdot E_{AB}(T-\tau)d\tau. \tag{6.4}$$

Transforming according to the construction in [Feller2] gives

$$\bar{E}_{BB}(s) = \frac{.15}{(s+.3)(s+.6)}.$$

By (2.3), we know that

$$\bar{G}_{BB}(s) = \frac{\bar{E}_{BB}}{1-\bar{E}_{BB}} = \frac{.15}{s^2+.9s+.03}.$$

Applying (2.3) also yields

$$\bar{G}_{BB}(s) = \bar{E}_{AB} + \bar{E}_{AB} \cdot \bar{G}_{BB} = \frac{.3s+.18}{s^2+.9s+.03}.$$

Next, note that S_B is the "reliability" function of state B, given that the state was B at $T = 0$. Thus $S_B(t) = e^{-.6t}$, so by (2.4), we obtain

$$\bar{P}_{AB}(s) = \frac{s+.3}{s^2+.9s+.03}.$$

From model II, we require G_{XX} and P_{XX}. We take $r = 2$ and $s = 3$ which are of course intended to have didactic value if not realism. In a manner similar to that made in the computation for model I is obtained

$$\bar{E}_{XX}(s) = \frac{6}{(s+3)(s+2)}$$

$$\bar{G}_{XX}(s) = \frac{6}{s^2 + 5s}. \tag{6.5}$$

Then we have from (2.5),

$$\bar{P}_{XX}(s) = \frac{1}{s+2}\left(1 + \frac{6}{s^2+5s}\right) = \frac{s+3}{s^2+5s}.$$

Now define $H_B(t) = G_{BB}(t) \cdot P_{XX}(t) + P_{BB}(t) \cdot G_{XX}(t)$. We are in fact interested in $\bar{H}_B(s)$. To find this one must invert the transforms $\bar{G}_{BB}(s)$, $\bar{P}_{XX}(s)$ and so on, perform multiplication and addition in the *time* domain, and then re-transform. A numerical mathematics package is helpful here. By this means one obtains

$$\bar{H}_B(s) = \frac{\sum_{i=1}^{3} u_i s^{3-i}}{\sum_{i=1}^{5} v_i s^{5-i}},$$

where

$$\vec{u} = [6.1500 \quad 34.6350 \quad 13.0995]$$
$$\vec{v} = [1.0000 \quad 11.8000 \quad 39.3700 \quad 26.9040 \quad 0.8859].$$

By (4.7) we have

$$\bar{Z}_{BB,XX}(s) = \bar{H}_B(s) \cdot \left[\frac{1}{s} - \bar{Z}_{BB,XX}\right].$$

Writing $\bar{Z}_{BB,XX}$ in rational form as $Z_B^{top}(s)/Z_B^{bot}(s)$ yields

$$Z_B^{top} = H_B^{top}, \quad Z_B^{bot} = s \cdot \left(H_B^{top} + H_B^{bot}\right), \tag{6.5}$$

giving $Z_B = \sum_{i=1}^{6} w_i s^{6-i}$.

Here $\vec{w} = [1.0000 \quad 11.8000 \quad 45.5200 \quad 61.5390 \quad 13.9854 \quad 0.00]$.

Next we require $H_A(t) = G_{AB}(t) \cdot P_{XX}(t) + P_{AB}(t) \cdot G_{XX}(t)$. Setting $\bar{H}_A(s) = H_A^{top}/H_A^{bot}$, we find that

$$H_A^{top} = \sum_{i=1}^{4} u_i s^{4-i}, \quad H_A = \sum_{i=1}^{5} v_i s^{5-i}, \tag{6.6}$$

where

$$\vec{u} = [0.3000 \quad 2.8500 \quad 10.7010 \quad 13.8294]$$
$$\vec{v} = [1.0000 \quad 11.8000 \quad 39.3700 \quad 26.9040 \quad 0.8859].$$

Next we apply (4.8) and obtain using a similar notation

$$Z_A^{top}(s) = H_A^{top}\left(Z_B^{bot} - s \cdot Z_B^{top}\right)$$
$$Z_A^{bot}(s) = s \cdot H_A^{bot} Z_B^{bot}. \tag{6.7}$$

The numerator of $\bar{Z}_A$ has degree 7 in s, and the denominator has degree 9, but they have 5 roots in common. When the corresponding factors have been canceled, what remains is

$$\bar{Z}_{AB,XX}(s) = \frac{\sum_{i=1}^{3} a_i s^{3-i}}{\sum_{i=1}^{5} b_i s^{5-i}},$$

where

$$\vec{a} = [1.0000 \quad 6.9000 \quad 17.7300]$$
$$\vec{b} = [1.0000 \quad 9.2000 \quad 21.6000 \quad 5.3790 \quad 0.0000].$$

Now $\bar{Z}_{AB,XX}(s)$ has distinct poles, and its partial fraction expansion corresponds to the explicit exponomial form of $Z_{AB,XX}(t)$. Writing simply $\bar{Z}$ and Z, we have

$$\bar{Z}(s) = \sum_{i=1}^{4} \frac{\sigma_i}{s + \rho_i},$$

where we write $\vec{\sigma}$ and $\vec{\rho}$ in column form

$$\vec{\sigma} = \begin{array}{l} \\ 0.98884551031790 \\ -0.05848988319612 \\ 0.08378848442687 \\ -1.01414411154865 \end{array}$$
$$\vec{\rho} = \begin{array}{l} \\ 0 \\ -5.35172855471732 \\ -3.56645190997164 \\ -0.28181953531104 \end{array}$$

Then of course $Z(t) = \sum_{i=1}^{4} \sigma_i e^{\rho_i t}$.

Consider the SHARPE input file for the combined model, together with the output information about node 4 in Figure 6-3. The distribution given is conditional upon entering the absorbing state 4. When multiplied by the given entrance probability, this gives the unconditional distribution, which is seen to agree with $Z_{AB,XX}(t)$ to 9 digits of accuracy.

```
bind
alph .3
bet .5
lam .1
r 2.
s 3.
end
markov death
1 4 alph
1 2 r
2 1 s
2 3 alph
3 2 bet
3 4 s
3 5 lam
end
1 1.
end
cdf(death,4)
end
```

```
information about system death node    4

probability of entering node: 9.88845510e-01

conditional CDF for time of reaching this absorbing state

   1.00000000e+00 t( 0) exp( 0.00000000e+00 t)
- 1.02558398e+00 t( 0) exp(-2.81819535e-01 t)
+ 8.47336450e-02 t( 0) exp(-3.56645191e+00 t)
- 5.91496676e-02 t( 0) exp(-5.35172855e+00 t)

mean:    3.62644539e+00
variance: 1.26658132e+01
```

Figure 6-3

In the second example we depict several physical components and their failure modes hierarchically. New features which were not present in the first example include

(1) determination of a simple exponential form of a distribution given its mean and variance,
(2) semi-Markov transitions,
(3) double poles in certain transition transforms,
(4) trigonometric solutions,
(5) neither coincident state is an initial state.

Duplex Node System

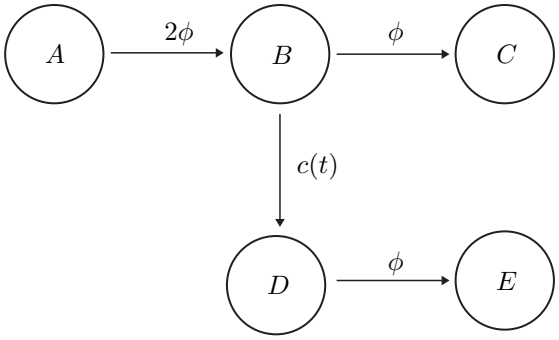


Figure 6-4

The example is a simplification of one aspect of the Integrated Airframe/Propul sion Control System Architecture (IAPSA). See [Cohen et al], p. 71. The *nodes* (sensor-processor pairs) form a reconfigurable duplex. The failure rate of each component is $\phi = .003$, the resulting model is shown in Figure 6-4. Here C and E are failure states, but as in the previous example we are concerned with failure modes arising from coincident conditions on separate structural levels. The transition function $c(t)$ represents the distribution of system reconfiguration time. It is the distribution of the random variable which is the sum of the times taken by the duplex operating system to detect an error, isolate the faulty unit, and configure to a simplex system. Experimentation with faults injected into the system has yielded a mean time of .01 sec with a variance of .001 sec^2. According to section 5, a hyper-exponomial distribution can be used for $c(t)$. A SHARPE model, and output realizing this are given in Figure 6-5, model "reconfig".

The other hierarchical component of the system is a dual partition network to which the nodes are attached. For simplicity we assume that either of two states can hold: both partitions are functioning, or else one partition is functioning and the other is undergoing repair (by configuring in a spare communication link). The "degraded" network is fully functional when the "node" system I is in either a stable duplex or simplex mode. However, the overall system cannot tolerate a simultaneous partition repair and duplex-to-simplex reconfiguration. The two-state model in Figure 6-6 illustrates the communication network, model II.

```
bind
p .025
mu 7.
lam 150
end
markov reconfig
1 3 mu
2 3 lam
end
1 p
2 1.-p
end
cdf(reconfig)
end
```

```
CDF for system reconfig:


   1.00000000e+00 t( 0) exp( 0.00000000e+00 t)
- 2.50000000e-02 t( 0) exp(-7.00000000e+00 t)
- 9.75000000e-01 t( 0) exp(-1.50000000e+02 t)


mean:    1.00714286e-02
variance: 1.00564116e-03
```

Figure 6-5

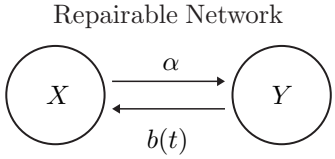


Figure 6-6

The partition failure rate is taken as a constant $\alpha = .01$; due to a rather complete understanding of the link repair mechanism, the repair distribution $b(t)$ is precisely known and is shown in Figure 6-7 (model net-repair). As indicated in the SHARPE output, the mean and variance of repair are roughly .02 sec. and .0003 sec^2 respectively. Note the factor of t in one of the terms of $b(t)$.

In the notation of the last example we are concerned with the function $Z_{AB,XY}(T)$. This is the probability given that model I begins (at $t = 0$) in A and model II begins in X, that *before the time* $t = T$ model I has been in B simultaneous with model II being in state Y. to this end one must find, for model I, the quantities G_{BB}, P_{BB}, G_{AB}, and P_{AB}. For model II, one seeks G_{YY}, P_{YY}, G_{XY}, and P_{XY}.

```
bind
p .02
mu 16.
lam 100
end
markov net-repair
1 2 mu
3 4 lam
4 5 lam
end
1 p
3 1.-p
end
cdf(net-repair)
end
```

```
CDF for system net-repair:


+ 1.00000000e+00 t( 0) exp( 0.00000000e+00 t)
- 9.80000000e+01 t( 1) exp(-1.00000000e+02 t)
- 9.80000000e-01 t( 0) exp(-1.00000000e+02 t)
- 2.00000000e-02 t( 0) exp(-1.60000000e+01 t)


mean:    2.08500000e-02
variance: 3.09527500e-04
```

Figure 6-7

Since B is not a recurrent state, we immediately obtain $G_{BB} = 0$ and thus $P_{BB}(T)) = 1 - S_B(T)$ from (2.4), second equation. A calculation of the distributions F_{BC} and F_{BD} yields

$$S_B = F_{BC} + F_{BD} = 1 - pe^{-(\phi+\mu)t} - qe^{-(\phi+\lambda)t}.$$

Since $G_{AB} = E_{AB} = dF_{AB} = 2\phi e^{-2\phi t}$, we also have

$$P_{AB}(T) = \int_0^T 2\phi e^{-2\phi\tau} \left[1 - S(T-\tau)\right] d\tau.$$

We could also obtain P_{AB} directly from the SHARPE model in Figure 6-8. The quantity P_{AA} is obviously $e^{-2\phi t}$, and the SHARPE output gives the total failure distribution, that is $P_{AC} + P_{AD}$, so P_{AB} is 1 minus the sum of these two quantities.

Note how in the model, the semi-Markov transition is entered as a general distribution.

```
bind
phi .003
pC .025
qC .975
mu 7.
lam 150.
end
semimark nodes
1 2 exp(2*phi)
2 3 exp(phi)
2 4 genN
1,0,0N
-pC,0,-muN
-qC,0,-lam
end
1 1.
end
cdf(nodes)
end
```

```
CDF for system nodes:

   1.00000000e+00 t( 0) exp( 0.00000000e+00 t)
- 1.00006044e+00 t( 0) exp(-6.00000000e-03 t)
+ 2.14377590e-05 t( 0) exp(-7.00300000e+00 t)
+ 3.90007800e-05 t( 0) exp(-1.50003000e+02 t)
```

Figure 6-8

Converting to the s-domain, one has $\bar{G}_{AB}(s) = 2\phi/(s+2\phi)$ and

$$\bar{P}_{AB}(s) = \frac{2\phi}{s+2\phi}\left[\frac{p}{s+\phi+\mu} + \frac{q}{s+\phi+\lambda}\right].$$

Also, $\bar{G}_{YY} = \bar{E}_{YY} + \bar{E}_{YY} \cdot \bar{G}_{YY}$ from (2.3). In fact $\bar{G}_{YY}(s) = \bar{G}^{top}_{YY}/\bar{G}^{bot}_{YY}$ where

$$\begin{aligned}\bar{G}^{top}_{YY} &= 10^3 \times (0.00000320s^2 + 0.09864000s + 1.6000)\\ \bar{G}^{bot}_{YY} &= 10^5 \times (0.00001000s^4 + 0.0021601000s^3 + 0.13202156800s^2 +\\ &\quad + 1.60033360000s + 0).\end{aligned}$$

Next, from (2.3), second equation, we have

$$\bar{G}_{XY}(T) = \bar{E}_{XY}(T) + \int_0^T \bar{E}_{XY}(\tau) G_{YY}(T-\tau)\, d\tau,$$

so

$$\bar{G}_{XY}(s) = \frac{.01}{s+.01}\left[\frac{\bar{G}^{bot}_{YY} + \bar{G}^{top}_{YY}}{\bar{G}^{bot}_{YY}}\right].$$

Next, from (2.4), letting L_s denote Laplace transform

$$\bar{P}_{XY}(s) = \bar{G}_{XY}(s) \cdot L_s\left[1 - b(t)\right] = \bar{G}_{XY}(x) \cdot \left\{\frac{100V}{(s+100)^2} + \frac{V}{s+100} + \frac{1-V}{s+16}\right\}.$$

Here $V = .98$ as indicated in Figure 6-7 (model repair-net).

We begin computing the quantities that govern the coincident states. Firstly, since $G_{BB}(t) = 0$ we have from (4.7)

$$Z_{BB,YY}(T) = \int_0^T H_{BY}(t) \cdot \left[1 - Z_{BB,YY}(t)\right] dt,$$

where $H_{BY}(t) = P_{BB}(t) \cdot G_{YY}(t)$. Solving yields

$$\bar{Z}_{BB,YY} = \frac{\bar{H}^{top}_{BY}}{s\left(\bar{H}^{bot}_{BY} + \bar{H}^{top}_{BY}\right)}.$$

After some simplification, one arrives at

$$\bar{Z}_{BB,YY}(s) = \sum_{j=1}^{10} \frac{\gamma_j}{s+\delta_j},$$

$\vec{\gamma} = 10^{-5}\times$	$\vec{\delta} =$
4.1360951	0
$2.6757974 - 1220.9445298i$	$-2.500031 + 0.001565i$
$2.6757974 + 1220.9445298i$	$-2.500031 - 0.001565i$
0.1175371	-1.660038
-6.5008554	-1.500127
$0.2214496 - 11.7051909i$	$-1.070078 + 0.009775i$
$0.2214496 + 11.7051909i$	$-1.070078 - 0.009775i$
0.0217117	-0.230031
-3.5689825	-0.070032

Manipulation of (4.8) yields the formulas

$$\bar{Z}^{top}_{AB,XY} = \bar{H}^{top}_{AX} \cdot \left(\bar{Z}^{bot}_{BY} - s\bar{Z}^{top}_{BY}\right)$$
$$\bar{Z}^{bot}_{AB,XY} = \bar{H}^{bot}_{AX} \cdot \bar{Z}^{bot}_{BY}.$$

Here $H_{AX}(t) = G_{AB}(t)P_{XY}(t) + P_{AB}(t)G_{XY}(t)$. We present H_{AX} explicitly; its value when $s = 0$ is of interest in that it represents the long-term or steady-state

arrival density. Since in practice $Z_{BB,YY}(t)$ is very small, formula (4.8) shows that the long-term probability of ending up in our coincident failure state (B, Y), instead of one of the other failure states, should be very close to this number which is 3.09×10^{-4}. Approximating our semi-Markov models by constant rate models gives an estimate of 2.99×10^{-4} for this probability. We do not give $Z_{AB,XY}(t)$ explicitly, since there are many terms. There is a strong temptation to simplify $\bar{Z}(s)$ by canceling roots in numerator and denominator which seem equal or very close, but this is a numerically delicate procedure. Instead we give $H_{AX}(t)$ explicitly from the partial fraction expansion.

$$\bar{H}_{AX}(s) = \bar{H}_{AX}^{top}/\bar{H}_{AX}^{bot} = \sum_{j=1}^{10} \sigma_j/(s+\rho_j)$$

We obtain

$\vec{\sigma} = 10^{-6}\times$	$\vec{\rho} =$
$-0.000038 + 0.001930i$	$-2.500079 + 0.009899i$
$-0.000038 - 0.001930i$	$-2.500079 - 0.009899i$ -0.389927 -1.500030
-0.214333	-0.070030
$-0.596018 + 29.687109i$	$-1.000109 + 0.009899i$
$-0.596018 - 29.687109i$	$-1.000109 - 0.009899i$
-0.074953	-0.160062
1.854995	-0.000060
0.016351	-1.000063

Then writing

$$\bar{H}_{AX}^{top} = \sum_{i=1}^{8} a_i s^{i-1}, \quad \bar{H}_{AX}^{bot} = \sum_{i=1}^{10} b_i s^{i-1},$$

we get

$\vec{b} =$	$\vec{a} =$
0	$1.0000e+00$
$1.2000e-049$	$7306e+02$
$1.0796e-013$	$8452e+05$
$3.8301e+017$	$9588e+07$
$6.7934e+039$	$2674e+09$
$6.2567e+056$	$0156e+11$
$2.7433e+071$	$9872e+13$
$4.2901e+082$	$6289e+14$
$1.9499e+091$	$0529e+15$

Letting $\sigma_j = d_j + e_j i$, $\rho_j = u_j + v_j i$, where $i = \sqrt{(-1)}$, results in

(7.7)
$$H_{AX}(t) = \sigma_3 e^{\rho_3 t} + \sigma_6 e^{\rho_6 t} + \sigma_7 e^{\rho_7 t} + \sigma_8 e^{\rho_8 t} + \sigma_9 e^{\rho_9 t} + \\ + 2e^{u_1 t}(d_1 \cos v_1 t + e_1 \sin v_1 t) + 2e^{u_4 t}(d_4 \cos v_4 t + e_4 \sin v_4 t).$$

7. Conclusions

An important recent approach in reliability (and performance) theory is found in the use of closed-form, analytical solutions. One advantage is that this approach lends itself very well to models which are built up of smaller submodels in a hierarchical fashion. In this manner fault arrival behavior, system response, architectural fault-tolerance features, and operating system features can be analyzed separately. Each model yields an analytic expression, which can then be put together according to formulas valid for the underlying stochastic process.

In practice, closed-form hierarchical solution of dependability problems has seen limited use. One limitation is that in combining two models, new failure states may have to be considered, which do not arise naturally from any particular failure state of either constituent submodel. We have presented a method for resolving such a situation. Using our formulas, it would seem feasible to incorporate the possibility of failure arising from the interaction of different hierarchical levels into a solution package such as SHARPE. The point of view we have presented emphasizes certain density functions and distributions arising in the study of semi-Markov processes. These quantities shed new light even on constant-rate processes, and are the key to solving models by decomposition. Large classes of (cyclic) semi-Markov chains can now be solved using the foundations laid in this article. The question of the numerical robustness of the closed-form approach is still an open one. This does not detract from the fact that "exponomial" methods are of great potential value in solving the problems of reliability modeling, which remain of both practical and theoretical interest.

References

Gayle Baker, Charles Stark Draper Laboratories, Cambridge, Mass, *personal communication*, [Baker].

S. Bavuso and P. Peterson, *CARE III Model Overview and User's Guide* (1985), no. 86404, NASA Tech. Memorandum, [Bavuso & Peterson].

T.A. Burton, *Volterra Integral and Differential Equations*, New York, Academic Press, 1983, [Burton].

R.W. Butler, An Abstract Language for Specifying Markov Reliability Models, IEEE Transactions on Reliability **R-35** (December, 1986), no. 5, [Butler].

R. Butler and A. White, *SURE Reliability Analysis, Program and Mathematics* (March, 1988), no. 2764, NASA Tech. Paper, [Butler & White].

R.V. Churchill, *The Inversion of the Laplace Transformation by a Direct Expansion in Series*, Math. Zeitschrift **42** (1937), 567–579, [Churchill].

G. Cohen, C. Lee. L. Brock and J. Allen, *Design/Validation Concept for an Integrated Airframe/ Propulsion Control System Architecture (IAPSA II)*, no. 178084, NASA Contractor Report, [Cohen et al].

E. Ellis and R. Butler, *Estimating the Distribution of Fault Latency in a Digital Processor*, NASA Tech. Memorandum (November, 1987), no. 0521, Langley Res. Center, [Ellis & Butler].

W. Feller, *An Introduction to Probability Theory and Its Applications* Vol. II, John Wiley New York, 1971, [Feller1].

W. Feller, *On the Integral Equation of Renewal Theory*, Annals of Math. Stat. **12** (1941), 243–267, [Feller2].

W. Feller, *On Semi-Markov Processes*, Proc. Natl. Acad. Sci. **51** (1964), 653–659, [Feller3].

G.B. Finelli, *Characterization of Fault Recovery through Fault Injection of FTMP*, IEEE Trans. Reliabilit **R-36** (June, 1987), no. 2, 164–170, [Finelli].

G. Golub and C. Van Loan, *Matrix Computations*, Johns Hopkins Univ. Press, Baltimore, 1983, [Golub & Van Loan].

J. Kohlas, *Stochastic Methods of Operations Research* (1982), Cambridge University Press, [Kohlas].

P.K. Lala, *Fault Tolerant & Fault Testable Hardware Design*, Prentice-Hall Englewood Cliffs, 1985, [Lala].

A.J. Lotka, *"A Contribution to the Theory of Self-renewing Aggregates, with Special Reference to Industrial Replacement"*, Annals of Math. Stat. (1939), 1–25, [Lotka].

A. Reibman and K. Trivedi, *Numerical Transient Analysis of Markov Models*, Computers and Operations Research **15** (1988), no. 1, 19–36, [Reibman & Trivedi].

S.M. Ross, *Applied Probability Models with Optimization Applications*, Holden-Day, San Francisco, 1983, [Ross].

P. Protzel, *Problems with Decomposition*, *NASA Langley Res. Center* Institute for Computer Applications in Science and Engineering (November, 1987), [Protzel].

R.A. Sahner and K.S. Trivedi, *A Hierarchical, Combinatorial-Markov Method of Solving Complex Reliability Models*, Fall Joint Computer Conference Proceedings (November, 1986), Dallas, [Sahner & Trivedi1].

R.A. Sahner and K.S. Trivedi, *"SHARPE: Symbolic Hierarchical Automated Reliability and Performance Evaluator" Guide for Users*, Duke Univ. Dept of Computer Science (September, 1986), [Sahner & Trivedi2].

J.A. Sjogren, *Closed-form Solution of Decomposable Stochastic Models*, Computers and Mathematics with Applications **23** (1992), no. 2, 43–68, [Sjogren].

K.S. Trivedi, *Probability & Statistics with Reliability, Queuing, and Computer Science Applications* 2nd Edition, John Wiley & Sons, New York, 2002, [Trivedi].

K. Trivedi, J. Dugan, R. Geist and M. Smotherman, *"Modeling Imperfect Coverage in Fault-Tolerant Systems"*, FTCS-14 *Digest* **84CH2050-3** (1984), 77–82, [Trivedi et al].

J.H. Wilkinson, *The Algebraic Eigenvalue Problem* 1st Edition, Oxford University Press, 1988, [Wilkinson].