APPENDIX

MARKOV RENEWAL PROCESSES

A.1. Introduction

Throughout this book we apply Markov Renewal Processes (MRP's) as one of the mathematical tools of analysis. In the appendix we briefly sketch the preliminary properties of renewal processes and MRP's.

We discuss the failure rate of the lifetime distribution of an item, where an item is referred to as a part, an equipment, a material, and so on. The probability law of the lifetime of an item is given by a distribution with the non-negative random variable X:

$$F(t) = \Pr\{X \leq t\}. \tag{A.1}$$

The probability density function is given by

$$f(t) = dF(t)/dt\ , \tag{A.2}$$

when it exists (we restrict ourselves to the continuous random variables throughout this book). Then, we define the <u>failure rate</u>

$$r(t) = f(t)/\overline{F}(t)\ , \tag{A.3}$$

when it exists and $\overline{F}(t) \equiv 1 - F(t) > 0$. We can explain the failure rate as follows: $r(t)dt$ is the probability that an item of age t will fail in the interval $(t, t+dt]$, given that the item survived at age t.

If we assume an exponential lifetime distribution

$$F(t) = 1 - e^{-\lambda t} \qquad (t \geq 0), \tag{A.4}$$

then

$$r(t) = \lambda e^{-\lambda t}/e^{-\lambda t} = \lambda\ , \tag{A.5}$$

which is constant. This property implies that a used exponential item is essentially "as good as new." Only the exponential distribution satisfies this property.

We discuss the exponential distribution from the viewpoint of the conditional survival probability. Consider the conditional survival probability

$$\Pr\{X > t + x | X > t\} = \Pr\{X > t + x\}/\Pr\{X > t\} , \qquad (A.6)$$

which is the conditional probability that an item survives at time $t + x$, given that it survived at time t. If we assume the exponential distribution in (A.4), then

$$\Pr\{X > t + x | X > t\} = e^{-\lambda(t+x)}/e^{-\lambda t} = e^{-\lambda x} , \qquad (A.7)$$

which is independent of the age history of the item. That is, the exponential distribution is independent of the age history of the item (i.e., how long the item is used since of the fresh item). We call this property as the <u>memoryless property.</u> In general, the memoryless property is given by the conditional survival probability

$$\Pr\{X > t + x | X > t\} = \Pr\{X > x\} , \qquad (A.8)$$

or

$$\bar{F}(t + x) = \bar{F}(t)\bar{F}(x) , \qquad (A.9)$$

by using the survival probability $\bar{F}(t) \equiv 1 - F(t)$. Let $\bar{F}(t)$ of a nondegenerate non-negative random variable satisfy (A.9) for all $t \geq 0$, $x \geq 0$. Then, we can show that $F(t)$ is an exponential distribution for some $\lambda > 0$ in (A.4). That is, the exponential distribution is necessary and sufficient for the functional equation (A.9) (see Barlow and Proschan (1975), Chapter 3).

If we assume a non-exponential distribution, the memoryless property is not satisfied any more. That is, we should know how long the item is used since of the fresh item to analyze the stochastic behavior. For instance, we assume a gamma (or Erlang) distribution

$$F(t) = 1 - (1 + 2\lambda t)e^{-2\lambda t} , \qquad (A.10)$$

where the mean is $1/\lambda$. Then, the failure rate is given by

$$r(t) = [dF(t)/dt]/\bar{F}(t) = 4\lambda^2 t/(1 + 2\lambda t) , \quad (A.11)$$

which is increasing in t. That is, we never satisfy the memoryless property for this case.

A.2. Renewal Processes

A renewal process is defined as a sequence of independent, non-negative, and identically distributed random variables X_1, X_2, ..., which are not degenerate at time 0. Let us consider an example of a renewal process. For instance, if we consider a replacement problem of identical lamps during an infinite time operation, this replacement problem is described by a renewal process, where X_i is the lifetime of a lamp.

Let $F(t)$ denote the inter-arrival distribution of the random variable X_i $(i = 1, 2, \ldots)$. Define the random variable $N(t)$ as the number of renewals (replacements) in $(0, t]$. Then

$$\begin{aligned} \Pr\{N(t) = n\} &= \Pr\{X_1 + X_2 + \ldots + X_n \le t \text{ and} \\ &\qquad X_1 + X_2 + \ldots + X_{n+1} > t\} \\ &= \Pr\{X_1 + X_2 + \ldots + X_n \le t \\ &\quad - \Pr\{X_1 + X_2 + \ldots + X_{n+1} \le t\} \\ &= F^{(n)}(t) - F^{(n+1)}(t) , \end{aligned} \quad (A.12)$$

where $F^{(n)}(t)$ is the n-fold Stieltjes convolution of $F(t)$ with itself, and $F^{(0)}(t)$ is a unit step function at $t = 0$. Let $M(t)$ denote the renewal function which is the expected number of renewals in $(0, t]$. Then

$$\begin{aligned} M(t) &= E[N(t)] \\ &= \sum_{k=0}^{\infty} k \Pr\{N(t) = k\} \\ &= \sum_{k=1}^{\infty} F^{(k)}(t) . \end{aligned} \quad (A.13)$$

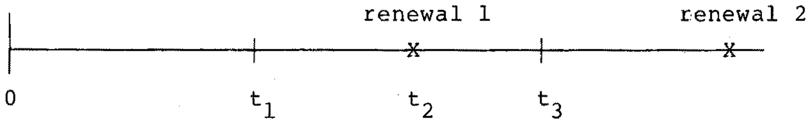

Fig. A.1. A sample function of a renewal process.

We assume the non-exponential distribution $F(t)$ in a renewal process and consider a sample function shown in Fig. A.1. The time instant 0 is independent of the history since it is a starting point. However, the time instant t_1 (or t_3) is not independent of the history, but depends on the time duration t_1 (or $t_3 - t_2$) since the non-exponential distribution has not the memoryless property. The time instant t_2 is independent of the history since it restarts as the random variable X_2. From Fig. A.1, we define that such the time instants 0 and t_2 are regeneration points and such the time instants t_1 and t_3 are non-regeneration points. Renewal processes are developed on the basis of the regeneration points (or the regenerative phenomena).

If we assume the exponential distribution in (A.4) in a renewal process, then we can show that any time instant is a regeneration point. That is, we can specify at any time instant as a starting point and obtain the corresponding results without any difficulty (recall the memoryless property of the exponential distribution). In general, a renewal process with the exponential inter-arrival distributions is called a Poisson process (see Ross (1970)).

A.3. Markov Renewal Processes

We are interested in only the number of renewals in a renewal process. That is, a renewal process is a one-state process which revisits (renews) infinitely often during an infinite time duration.

An MRP is a stochastic process in which some different states are introduced, the transition probabilities and the inter-arrival distributions from one state to another are specified. We shall define an MRP and briefly give some interesting results. We restrict ourselves to an MRP with finitely many states since the applications in reliability theory are mainly described by an MRP with finitely many states.

An MRP with finitely many states is defined as follows: Assume a finite number of states $i = 0, 1, 2, \ldots, N$. Define the transition probability

$$Q_{ij}(t) = \Pr\{\text{after making a transition into state } i\text{, the process next makes a transition into state } j\text{, in an amount of time less than or equal to } t\},$$

for any i and j, where

$$Q_{ij}(0) = 0 \qquad (i, j = 0, 1, \ldots, N), \tag{A.14}$$

$$\sum_{j=0}^{N} Q_{ij}(\infty) = 1 \qquad (i = 0, 1, \ldots, N). \tag{A.15}$$

We must define that the time instant i $(i = 0, 1, 2, \ldots, N)$, at which the process just makes a transition into state i, is a regeneration point. We define the unconditional sojourn distribution in state i

$$H_i(t) = \sum_{j=0}^{N} Q_{ij}(t) \qquad (i = 0, 1, 2, \ldots, N), \tag{A.16}$$

not specifying any next visiting state. We define the random variable $X(t)$, where $X(t) = i$ denotes that the process is in state i at time t. We also define the random variable $N_i(t)$, where $N_i(t) = k$ denotes that the number of visit to state i is k in $(0, t]$. The Markov renewal process concerns with the random variables $N_i(t)$ $(i = 0, 1, \ldots, N)$. On the other hand, the semi-Markov process concerns with the random variable $X(t)$. We understand that both the MRP and the semi -Markov process are essentially the same stochastic processes. We just call the MRP's throughout this book.

Define the following quantities:

$$P_{ij}(t) = \Pr\{X(t) = j \mid X(0) = i\}, \tag{A.17}$$

$$G_{ij}(t) = \Pr\{N_j(t) > 0 \mid X(0) = i\}, \tag{A.18}$$

$$M_{ij}(t) = E[N_j(t) \mid X(0) = i], \tag{A.19}$$

for i and j $(i, j = 0, 1, 2, \ldots, N)$. We note that $P_{ij}(t)$ denotes the probability that the process is in state j at time t, given that it was in state i at time 0, $G_{ij}(t)$ the first-passage distribution from state i to state j in $(0, t]$, and $M_{ij}(t)$ the generalized renewal function in state j (i.e., the mean number of visit to state j in $(0, t])$, given that the process was in state i at time 0. Recall the Stieltjes convolution defined in (1.9). Applying the renewal-theoretic arguments, we have

$$P_{ii}(t) = 1 - H_i(t) + \sum_{k=0}^{N} Q_{ik}(t)*P_{ki}(t) , \tag{A.20}$$

$$P_{ij}(t) = \sum_{k=0}^{N} Q_{ik}(t)*P_{kj}(t) \qquad (i \neq j) , \tag{A.21}$$

$$G_{ij}(t) = Q_{ij}(t) + \sum_{\substack{k=0\\k\neq j}}^{N} Q_{ik}(t)*G_{kj}(t) , \tag{A.22}$$

$$M_{ij}(t) = G_{ij}(t) + G_{ij}(t)*M_{jj}(t) , \tag{A.23}$$

for i and j $(i, j = 0, 1, 2, \ldots, N)$.

Let $q_{ij}(s)$, $h_i(s)$, $p_{ij}(s)$, and $m_{ij}(s)$ denote the Laplace -Stieltjes (LS) transforms of $Q_{ij}(t)$, $H_i(t)$, $P_{ij}(t)$, $G_{ij}(t)$, and $M_{ij}(t)$, respectively. Let $\underline{q}(s)$ and $\underline{m}(s)$ denote the matrices composed of $q_{ij}(s)$ and $m_{ij}(s)$, respectively. Then we take the LS transforms (A.21), (A.22), and (A.23) and solve them:

$$\underline{m}(s) = [\underline{I} - \underline{q}(s)]^{-1}\underline{q}(s) = [\underline{I} - \underline{q}(s)]^{-1} - \underline{I} , \tag{A.24}$$

$$g_{ij}(s) = m_{ij}(s)/[1 + m_{jj}(s)] , \tag{A.25}$$

$$p_{jj}(s) = [1 - h_j(s)]/[1 - g_{jj}(s)] , \tag{A.26}$$

$$p_{ij}(s) = p_{jj}(s)g_{ij}(s) \qquad (i \neq j) , \tag{A.27}$$

for i and j $(i, j = 0, 1, 2, \ldots, N)$, where $\underline{I}$ is the identity matrix. Equations (A.24) - (A.27) tell us that $\underline{q}(s) = [q_{ij}(s)]$ implies the LS transforms $m_{ij}(s)$, $g_{ij}(s)$, and $p_{ij}(s)$. However, it is generally difficult to invert the LS transforms analytically except the simplest cases. We should apply the numerical inversion (see Bellman et al. (1966)).

It is very difficult to discuss the transient behavior analytically since all the results in (A.24) - (A.27) are given by the LS transforms. However, it might be quite easy to discuss the limiting behavior analytically. We first introduce the notion of an embedded Markov Chain (MC). That is, an embedded MC of an MRP is an MC in which any one-step transition duration from one state to another is regarded as a unit of time. The transition probability matrix $\underline{Q}$ of the embedded MC is derived by the limiting operation:

$$\underline{Q} = [q_{ij}] , \tag{A.28}$$

which we call as the limiting transition probability matrix since each element is the limiting transition probability

$$q_{ij} = \lim_{s \to 0} q_{ij}(s) = \lim_{t \to \infty} Q_{ij}(t) . \tag{A.29}$$

We assume that the embedded MC of the MRP is positive recurrent (i.e., all the states communicate). Then we have the positive stationary distribution $\underline{\pi} = [\pi_0 , \pi_1 , \ldots, \pi_N]$ as a unique solution to equations:

$$\underline{\pi} = \underline{\pi} \cdot \underline{Q} \qquad \text{and} \qquad \sum_{i=0}^{N} \pi_i = 1 \tag{A.30}$$

(see Ross (1970)). Barlow and Proschan (1965) showed the mean first passage ℓ_{ij} for the MRP as follows:

$$\ell_{ij} = \sum_{\substack{k=0 \\ k \neq j}}^{N} q_{ik} \ell_{kj} + \mu_i \qquad (i, j = 0, 1, \ldots, N), \tag{A.31}$$

where

$$\mu_i = \int_0^\infty t dH_i(t) \qquad (i = 0, 1, \ldots, N) \tag{A.32}$$

is the unconditional mean of the distribution $H_i(t)$. By using the stationary distribution $\underline{\pi}$ and the matrix manipulations, we have

$$\ell_{ii} = \sum_{k=0}^{N} \pi_k \mu_k / \pi_i \qquad (i = 0, 1, \ldots, N). \tag{A.33}$$

Generalizing the renewal theorem (see Ross (1970)), we can show

$$M_j \equiv \lim_{t \to \infty} M_{ij}(t)/t = 1/\ell_{jj} \qquad (i, j = 0, 1, \ldots, N). \tag{A.34}$$

Applying the above results to (A.26) and (A.27), we have

$$P_j \equiv \lim_{t\to\infty} P_{ij}(t) = \mu_j/\ell_{jj} = \pi_j\mu_j/\sum_{k=0}^{N} \pi_k\mu_k$$

$$(i, j = 0, 1, \ldots, N). \qquad (A.35)$$

Note that M_j and P_j are independent of a starting state i. The interesting results (A.33), (A.34), and (A.35) are based on the assumption that all the states of the embedded MC (or the MRP) communicate.

We have given the brief results of the MRP's. The detailed discussions were given by Pyke (1961a, 1961b), Barlow and Proschan (1965), Ross (1970), and Cinlar (1975). The unique modification of the regeneration point techniques of MRP's was developed by Nakagawa and Osaki (1974, 1976). Throughout this book we apply the unique modification of MRP's to the models of computer architectures to analyze the MRP's with some non-regeneration points.

REFERENCES

Some references listed below are not cited in the text. However, we include these references since they are closely related to the topics in this book. All references listed below are in alphabetical order.

Y. Abe (1977), "A Japanese On-Line Banking System," Datamation, vol. 23, pp. 89-97.

Y. Abe, Y. Ueno, A. Honma, E. Kato, Y. Nose, Y. Yasufuku and M. Atarashi (1977), "The Dai-Ichi Kangyo Bank, LTD - HOPS: Development and System Configuration," (in Japanese) FUJITSU, vol. 28, pp. 856-892.

T. Anderson and B. Randell (Ed.) (1979), Computing Systems Reliability, Cambridge University Press, Cambridge.

T.F. Arnold (1973), "The Concept of Coverage and Its Effect on the Reliability Model of a Repairable System," IEEE Trans. Comput., vol. C-22, pp. 251-254.

A. Avizienis (1976), "Fault-Tolerant Systems," IEEE Trans. Comput., vol. C-25, pp. 1304-1312.

A. Avizienis, G.C. Gilley, F.P. Mathur, D.A. Rennels, J.A. Rohr and D.K. Rubin (1971), "The STAR (Self-Testing And Repairing) Computer: An Investigation of the Theory and Practice of Fault-Tolerant Computer Design," IEEE Trans. Comput., vol. C-20, pp. 1312-1321.

R.E. Barlow and F. Proschan (1965), Mathematical Theory of Reliability, Wiley, New York.

R.E. Barlow and F. Proschan (1975), Statistical Theory of Reliability and Life Testing - Probability Models, Holt, Rinehart and Winston, New York.

M.D. Beaudry (1978), "Performance-Related Reliability Measures for Computing Systems," IEEE Trans. Comput., vol. C-27, pp. 540-547.

R. Bellman, R.E. Kalaba and J. Lockett (1966), Numerical Inversion of the Laplace Transform, American Elsevier, New York.

B.R. Borgerson and R.F. Freitas (1975), "A Reliability Model for Gracefully Degrading and Standby-Sparing Systems," IEEE Trans. Comput., vol. C-24, pp. 517-525.

W.G. Bouricius, W.C. Carter, D.C. Jessep, P.R. Schneider and A.B. Wadia (1971), "Reliability Modeling for Fault-Tolerant Computers," IEEE Trans. Comput., vol. C-20, pp. 1306-1311.

W.G. Bouricius, W.C. Carter and P.R. Schneider (1969), "Reliability Modeling Techniques for Self-Repairing Computer Systems," in Proc. 12th ACM Nat. Conf., pp. 295-309.

W.C. Carter and W.G. Bouricius (1971), "A Survey of Fault-Tolerant Computer Architecture and Its Evaluation," Computer, vol. 4, pp. 9-16.

G.A. Champine (1978), Computer Technology Impact on Management, North -Holland Publishing Company, Amsterdam.

T.C.K. Chou and J.A. Abraham (1980), "Performance/Availability Model of Shared Resource Multiprocessors," IEEE Trans. Reliab., vol. R-29, to appear.

D.K. Chow (1975), "Availability of Some Repairable Computer Systems," IEEE Trans. Reliab., vol. R-24, pp. 64-66.

E. Cinlar (1975), Introduction to Stochastic Processes, Prentice-Hall, Englewood Cliffs, New Jersey.

A.E. Cooper and W.T. Chow (1976), "Development of On-Board Computer Systems," IBM J. Res. Develop., vol 20, pp. 5-19.

A. Costes, C. Landrault and J.C. Laprie (1978), "Reliability and Availability Models for Maintained Systems Featuring Hardware Failures and Design Faults," IEEE Trans. Comput., vol. C-27, pp. 548-560.

A.L. Hopkins, Jr. and T.B. Smith (1975), "The Architectural Elements of a Symmetric Fault-Tolerant Multiprocessor," IEEE Trans. Comput., vol. C-24, pp. 498-505.

M. Kajiyama and S. Osaki (1979), "Performance-Related Analyses for Computer Systems with High Reliability," (in Japanese) Trans. IECE Japan, vol. J62-D, pp. 742-749.

M. Kinugasa and S. Osaki (1979), "Analysis of Highly Reliable Redundant Computer Systems Taking Account of Performance," (in Japanese) IECE Technical Reports on Reliability, R79-29, pp. 23-29.

M. Kinugasa and S. Osaki (1980), "Reliability Analysis of a Three-Unit Standby Redundant System," (in Japanese) Abstracts presented in the Spring Meeting of Operations Res. Soc. Japan, pp. 144-145.

I. Koren and S.Y.H. Su (1979), "Reliability Analysis of N-Modular Redundant Systems with Intermittent and Permanent Faults," IEEE Trans. Comput., vol. C-28, pp. 514-520.

J.C. Laprie (1976), "On Reliability Prediction of Repairable Redundant Digital Structures," IEEE Trans. Reliab., vol. R-25, pp. 275-277.

J. Losq (1976), "A Highly Efficient Redundancy Scheme: Self-Purging Redundancy," IEEE Trans. Comput., vol. C-25, pp. 569-578.

F.P. Mathur and A. Avizienis (1970), "Reliability Analysis and Architecture of a Highly Redundant Digital System: Generalized Triple Modular Redundancy with Self-Repair," in Proc. 1970 SJCC, AFIPS Conf. Proc., vol. 36, pp. 375-383.

F.P. Mathur and P.T. de Souza (1975), "Reliability Modeling and Analysis of Generalized Modular Redundant Systems," IEEE Trans. Reliab., vol. R-24, pp. 296-299.

T. Nakagawa and S. Osaki (1974), "Stochastic Behaviour of a Two-Unit Standby Redundant System," INFOR, vol. 12, pp. 66-70.

T. Nakagawa and S. Osaki (1976), "Markov Renewal Processes with Some Non-Regeneration Points and Their Applications to Reliability Theory," Microelectron. Reliab., vol. 15, pp. 633-636.

S. Osaki (1970), "System Reliability Analysis by Markov Renewal Processes," J. Operations Res. Soc. Japan, vol. 12, pp. 127-188.

S. Osaki (1974), "Signal-Flow Graphs in Reliability Theory," Microelectron. Reliab., vol. 13, pp. 539-541.

S. Osaki and T. Nakagawa (1976), "Bibliography for Reliability and Availability for Stochastic Systems," IEEE Trans. Reliab., vol. R-25, pp. 284-287.

S. Osaki and T. Nishio (1979), "Availability Evaluation of Redundant Computer Systems," Comput. Operations Res., vol. 6, pp. 87-97.

R. Pyke (1961a), "Markov Renewal Processes: Definitions and Preliminary Properties," Ann. Math. Statist., vol. 32, pp. 1231-1242.

R. Pyke (1961b), "Markov Renewal Processes with Finitely Many States," Ann. Math. Statist., vol. 32, pp. 1243-1259.

B. Randell (1975), "System Structure for Software Fault-Tolerance," IEEE Trans. Software Eng., vol. SE-1, pp. 220-232.

S.M. Ross (1970), Applied Probability Models with Optimization Applications, Holden-Day, San Francisco.

R.A. Short (1968), "The Attainment of Reliable Digital Systems Through the Use of Redundancy - A Survey," IEEE Comput. Group News, vol. 2, pp. 2-17.

Y. Suzuki, T. Nakagawa and Y. Sawa (1977), "Reliability Analysis of Computer Systems with High Reliability," (in Japanese) Trans. IECE Japan, vol. J60-D, pp. 1047-1052.

J. von Neumann (1956), "Probabilistic Logics and the Synthesis of Reliable Organizations from Unreliable Components," in Autamata Studies, C.E. Shannon and J. McCarthy (Ed.), Princeton University Press, Princeton, New Jersey.