Dagger-Sampling Monte Carlo For System Unavailability Evaluation

Hiromitsu Kumamoto
Kyoto University, Kyoto
Kazuo Tanaka
Kyoto University, Kyoto
Koichi Inoue, Member IEEE
Kyoto University, Kyoto
Ernest J. Henley
University of Houston, Houston

Key Words—**System unavailability, Monte Carlo method, Antithetic variates, Rare-event simulations.**

Reader Aids—
Purpose: Widen state of the art
Special math needed for explanations: Statistics, Boolean algebra
Special math needed for results: Statistics
Results useful to: Reliability and safety analysts

Abstract—**Reliability problems usually result in rare-event simulations, and hence direct Monte Carlo methods are extremely wasteful of computer time. This paper presents a new application of "dagger-sampling", for calculating the system unavailability of a large complicated system represented by a coherent fault tree. Since a small number of uniform random numbers generate a number of trials, dagger-sampling appreciably reduces computation time, and hence a large number of trials become possible for the rare-event problems. Further, dagger-sampling decreases the variance of the Monte Carlo estimator because it generates negatively correlated samples.**

1. INTRODUCTION

A Monte Carlo method consists of building, usually with a computer program, a probabilistic model of the system under investigation. The model is then run a large number of times, and each time the performance of the synthesized system is recorded.

Monte Carlo simulations are easy to carry out, and can be applied to systems that are too complex or too large to solve by deterministic methods such as KITT codes [1]. Most reliability problems, however, result in rare-event simulations, and hence direct Monte Carlo methods are extremely wasteful of computer time. If for example, the exact system unavailability is 10^{-5}, 10^4 trials could result in no system failure and might lead us to conclude that the system is completely available. On the average, we would require 10^5 trials to have a failure, and at least 10^7 trials are required to produce Monte Carlo estimates with one significant figure.

Refs. [2-10] apply Monte Carlo techniques to reliability problems. Some of them use direct sampling techniques, and others employ more sophisticated ones. For example, Easton & Wong [9] proposed a "sequential destruction method", embedding the original sampling space into a broader set. The present authors [10] restricted Monte Carlo trials to a subspace of the original sampling set, and constructed a "restricted-sampling Monte Carlo method". These two methods are theoretically guaranteed to reduce variance of the Monte Carlo estimator, while they deal with complex reliability problems.

The present paper gives a new application of "dagger-sampling" for calculating the system unavailability, given a *s*-coherent fault tree. Similar sampling techniques have been applied to queueing problems [11].

Along with reducing computation time, the dagger-sampling combines success and failure states, and generates negatively-correlated state vectors of basic events. This negative correlation at the basic-event level applies to the corresponding states of the top event since the system structure is a monotonically increasing function. The Monte Carlo estimator (the average top-event state) has a smaller variance than the direct Monte Carlo method because probabilistic fluctuations are cancelled out by the correlation. The variance reduction technique using the negative correlations is called "antithetic variates" in the field of Monte Carlo methods [11-13].

2. PROBLEM STATEMENT

2.1 Assumptions

1) The fault tree has k basic events, $1, \ldots, k$.
2) The basic events are *s*-independent.
3) The fault tree is *s*-coherent.
4) Every state vector is possible, i.e., for all X, the inequality

$$0 < \Pr(X) = \prod_{l=1}^{k} \Pr\{X_l\} < 1 \tag{2.1}$$

holds.
5) This is a rare-event reliability problem:

$$0 < P_l \leqslant ½ \text{ for some } l \tag{2.2}$$

2.2 Notation

X_l — state of basic event l,

$$X_l \equiv \begin{cases} 1, & \text{if event } l \text{ occurs} \\ 0, & \text{otherwise} \end{cases}$$

X — $(X_1, \ldots, X_k)$ is a basic event state vector.

$\psi(X)$ — binary function expressing the top event of the fault tree,

$$\psi(X) \equiv \begin{cases} 1, & \text{if the top event occurs} \\ 0, & \text{otherwise} \end{cases}$$

P_l probability of occurrence of basic event l.
N sample size of Monte Carlo methods.
z_i sample state vectors of dagger-sampling, $i = 1, \ldots, N$.
c_i sample state vectors of direct Monte Carlo, $i = 1, \ldots, N$.
Q_s exact system unavailability.
Q_D, Q_C system unavailability estimators by dagger-sampling, and by direct Monte Carlo.

We need to calculate the system unavailability

$$Q_s = \Pr\{\psi(X) = 1\} = \sum_X \psi(X)\Pr\{X\} \tag{2.3}$$

$$= E_X\{\psi(X)\}. \tag{2.4}$$

3. DAGGER-SAMPLING

In dagger-sampling, a small number of uniform random numbers generate several sample vectors which are negatively correlated. This can be best illustrated by simple examples.

Consider first the case where each basic event occurs with probability 0.01. Dagger-sampling generates 100 samples for event 1 in the manner shown in Fig. 1.

A group of 100 intervals between 0 and 1 is introduced for basic event 1. Interval i is used for generating event 1 in trial i, and has subinterval $[(i - 1) \times 0.01, i \times 0.01)$. The length of the subinterval is equal to 0.01, the probability of the occurrence of basic event 1. The first interval has subinterval $[0, 0.01)$, and the last interval $[0.99, 1)$.

Only one uniform random number is generated for the group of 100 unit-intervals. Assume that the random number falls in subinterval of interval i. Then, basic event 1 is assumed to occur in trial i and not to occur in the other 99 trials. For example, random number 0.4256 determines that event 1 occurs in trial 43 and does not occur in the other 99 trials (Fig. 1). The basic event occurs with probability 0.01 in each trial. One uniform random number pierces 100 intervals, and determines 100 trials for the basic event: hence the description "Dagger-sampling".

Basic event 2 is generated similarly by using another uniform random number for the group of 100 intervals. Any number of event-state vectors can be sampled by s-independent repetitions of each 100 generations.

The direct Monte Carlo method in this example generates 100 samples for event 1, using 100 uniform random numbers (Fig. 2). The dagger-sampling generates $100 \times m$ trials by $k \times m$ uniform random numbers, while the direct method generates m trials by the same $k \times m$ numbers. Thus, we need only 1/100 the random numbers of the direct Monte Carlo method. This means that the proposed method can generate 500 000 trials within the same computational effort as 5000 direct Monte Carlo trials.

We now consider the general cases where basic event l has probability P_l. Let $[1/P_l]$ be the largest integer not larger than $1/P_l$. Dagger-sampling generates $[1/P_l]$ samples for event l, using one random number in the manner shown in Fig. 3: we introduce $[1/P_l]$ subintervals; the length of each subinterval is P_l; if the random number is less than $[1/P_l]P_l$, then one out of $[1/P_l]$ samples is the occurrence of event l, similarly to Fig. 1; otherwise, all $[1/P_l]$ samples are the non-occurrence of event l (see trials 4-6 for event 1).

4. ESTIMATOR BASED ON DAGGER-SAMPLING

Let $z_1, \ldots, z_N$ be N event-state vectors generated by the dagger-sampling. The system unavailability Q_s of (2.4) can be estimated by the unbiased binominal estimator Q_D, the subscript D standing for "Dagger".

$$Q_D \equiv N^{-1} \sum_{i=1}^{N} \psi(z_i). \tag{4.1}$$

The variance of Q_D is the sum of variances of $\psi(z_i)$ and covariances of different $\psi(z_i)$ divided by N^2.

$$\mathrm{Var}\{Q_D\} = N^{-2}\Big(\sum_{i=1}^{N} \mathrm{Var}\{\psi(z_i)\} + \sum_{i\neq j} \mathrm{Cov}\{\psi(z_i), \psi(z_j)\}\Big) \tag{4.2}$$

The first sum of the r.h.s. of (4.2) is the variance of direct Monte Carlo estimator Q_C, since $\mathrm{Var}\{\psi(z_i)\} = \mathrm{Var}\{\psi(c_i)\}$ and $\mathrm{Cov}\{\psi(c_i), \psi(c_j)\} = 0$ for event-state vectors c_i and c_j generated by the direct Monte Carlo.

In dagger-sampling, two sample vectors z_i and z_j are not s-independent, but are correlated because a small number of uniform random numbers generate several sample vectors. Figs. 1 or 3 show that: if a basic event occurs during a trial, then the basic event does not occur in other trials for the same group. Thus, the correlation of two sample vectors z_i and z_j is negative as long as the vectors have some elements in the same group: if elements $z_{i,l}$ and $z_{j,l}$ in trials i and j are generated by a common random number, then

$$E\{z_{i,l}\, z_{j,l}\} - E\{z_{i,l}\}E\{z_{j,l}\} = -P_l^2 < 0.$$

Since the structure function $\psi(X)$ is s-coherent, it is monotonically increasing. Thus, the negative correlation between z_i and z_j also applies to $\psi(z_i)$ and $\psi(z_j)$, and we have, from (4.2) that

$$\mathrm{Var}\{Q_D\} < \mathrm{Var}\{Q_C\}. \tag{4.3}$$

As will be shown in the next section, most of the computation time for the direct or dagger-sampling Monte Carlo methods is spent in generating the uniform random numbers. Inequality (4.3) shows another characteristics of dagger-sampling; it yields a more accurate estimate for the system unavailability than the direct method, for a given sample size N.

A formal proof of (4.3) can be obtained, using the following theorem. This theorem itself can be proven easily by induction on l, using pivotal representation of ψ.

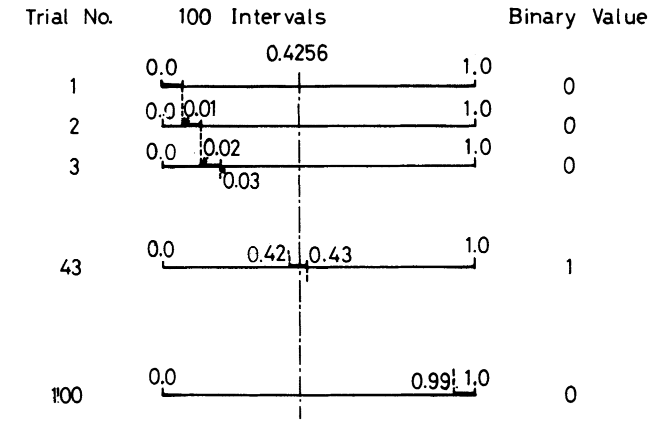

Fig. 1. Generation of 100 Samples for Basic Event 1. (Dagger-Sampling)

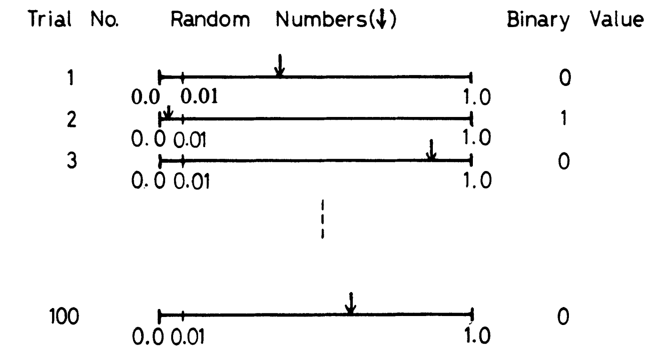

Fig. 2. Generation of 100 Samples for Basic Event 1. (Direct Monte Carlo)

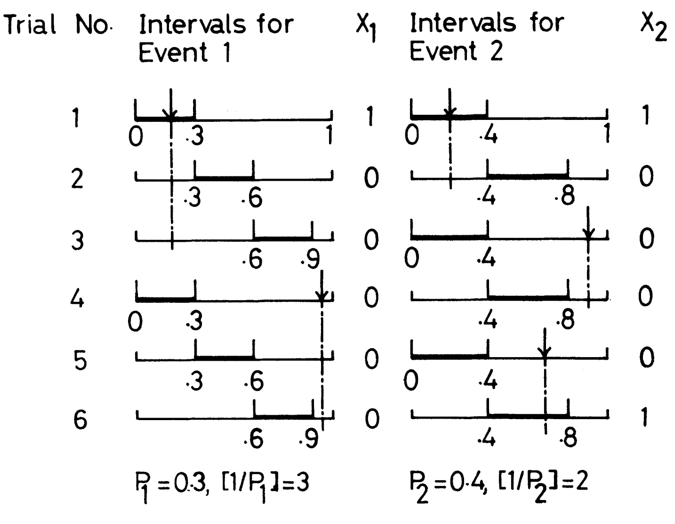

Fig. 3. Dagger-Sampling in General Case.

Theorem Assume the following conditions 1 to 6.

1) $\psi(X)$ is a s-coherent function of a k-dimension binary vector X.

2) $(x, y) = (x_1, \ldots, x_k, y_1, \ldots, y_k)$ is a $2k$-dimension binary-stochastic vector.

3) $\Pr(x, y) = \prod_{l=1}^{k} \Pr\{x_l, y_l\}.$ (4.4)

4) $0 < P_l \equiv \Pr\{x_l = 1\} = \Pr\{y_l = 1\} < 1.$ (4.5)

5) For each l, variables x_l and y_l are either negatively correlated or s-independent.

$$\Pr\{x_l = 1, y_l = 1\} \leq \Pr\{x_l = 1\}\Pr\{y_l = 1\} = P_l^2 \tag{4.6}$$

6) There exists an l such that a strong inequality (<) holds in (4.6).

Then, the covariance of $\psi(x)$ and $\psi(y)$ is negative.

$$\mathrm{Cov}\{\psi(x), \psi(y)\} < 0. \tag{4.7}$$

The fault tree satisfies Condition 1 by assumption. Let z_i an z_j be two sample vectors which have some elements in the same group. The existence of such vectors is guaranteed by (2.2). Conditions 2 to 6 hold for $x = z_i$ and $y = z_j$. Thus, from the theorem,

$$\mathrm{Cov}\{\psi(z_i), \psi(z_j)\} < 0. \tag{4.8}$$

If z_i and z_j belong to different groups, then these vectors are s-independent:

$$\mathrm{Cov}\{\psi(z_i), \psi(z_j)\} = 0. \tag{4.9}$$

Inequality of (4.3) is obtained by (4.2), (4.8), and (4.9).

5. CALCULATION OF ψ (SAMPLE VECTOR).

Most sample state vectors involve simultaneous occurrence of, at most, two basic events. For such state vectors, we can calculate the value of ψ by table-look-up.

$$\psi(\text{Sample Vector}) = \begin{cases} 1, & \text{if the vector corresponds to a cut set} \\ 0, & \text{otherwise.} \end{cases}$$

For a sample vector involving the occurrence of three or more basic events, ψ is calculated by simulating the fault tree for the state vector. This probability of occurrence is relatively small, and we can appreciably reduce the amount of computation required for the calculation of ψ. Thus, in most cases, the major computational effort consists of generating uniform random numbers.

6. A NUMERICAL EXAMPLE.

Consider the fault tree in Fig. 4 where each basic event occurs with probability 0.01. The termwise calculation of (2.3) gives the exact system unavailability as

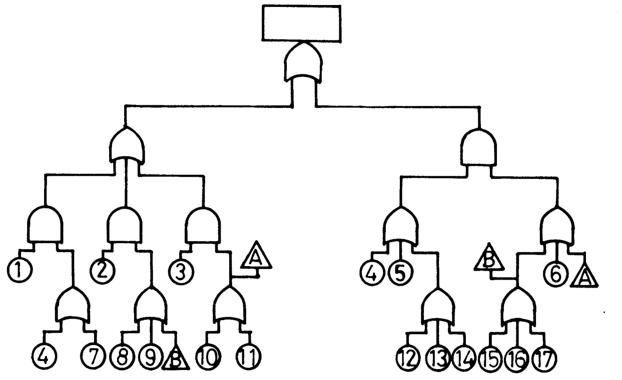

Fig. 4. Fault Tree for Numerical Example.

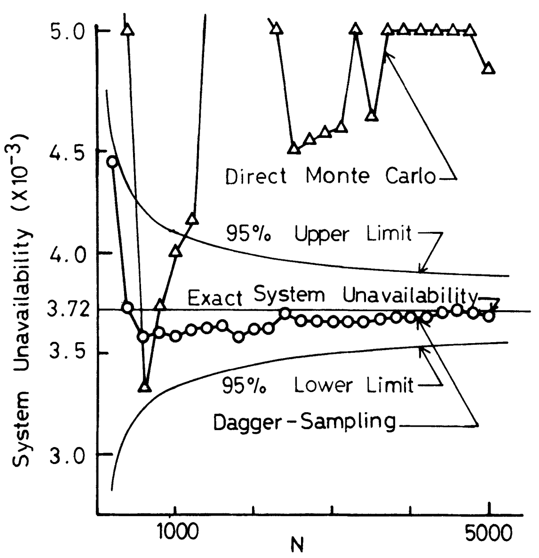

Fig. 5. Result of Direct and Dagger-Sampling Monte Carlo Method.

$Q_s = 3.72 \times 10^{-3}$.

Figure 5 shows the results of the dagger-sampling, and the direct Monte Carlo methods. The horizontal axis of the figure shows the number of trials of the direct Monte Carlo. For the present example, the dagger-sampling could generate 100 trials by the computation time for one direct Monte Carlo trial. Thus, Fig. 5 compares N direct Monte Carlo trials with $100 \times N$ dagger-sampling trials. Good estimates of the system unavailability are obtained by the dagger-sampling.

The exact variance of Q_D was obtained by termwise calculations of $\text{Var}\{\psi(z_i)\}$ and $\text{Cov}\{\psi(z_i)\psi(z_j)\}$ in (4.2). The s-confidence limit in Fig. 5 was calculated by $\text{Var}\{Q_D\}$, assuming asymptotic normality. A practical way for estimating $\text{Var}\{Q_D\}$ is to neglect the second term in (4.2) and to evaluate the first term by sample variance. This gives a conservative estimate for $\text{Var}\{Q_D\}$.

REFERENCES

[1] W.E. Vesely, R.E. Narum, "PREP and KITT: Computer codes for the automatic evaluation of a fault tree", Idaho Nuclear Corp., IN-1349, 1970*.

[2] S.J. Kamat, M.W. Riley, "Determination of reliability using event-based Monte Carlo simulation", *IEEE Trans. Reliability,* vol R-24, 1975 Apr, pp 73-75.

[3] S.J. Kamat, W.E. Franzmeier, "Determination of reliability using event-based Monte Carlo simulation Part II", *IEEE Trans. Reliability,* vol R-25, 1976 Oct, pp 254-255.

[4] P.W. Becker, "Finding the better of two similar designs by Monte Carlo techniques", *IEEE Trans. Reliability,* vol R-23, 1974 Oct, pp 242-246.

[5] M. Mazumdar, "Importance sampling in reliability estimation", *Reliability and Fault Tree Analysis,* Barlow, R.E. et al eds. SIAM, 1975, pp 153-163.

[6] R. Van. Slyke, H. Frank, "Network reliability analysis: Part I", *Networks,* vol 1, 1972, pp 279-290.

[7] R. Van. Slyke, H. Frank, A. Kershenbaum, "Network reliability analysis: Part II", *Reliability and Fault Tree Analysis,* Barlow, R. E. et al eds. SIAM, 1975, pp 619-650.

[8] F. Frank, I.T. Frish, "Analysis and design of survivable networks", *IEEE Trans. Commun. Technol.,* vol COM-18, 1970 Oct, pp 501-519.

[9] M.C. Easton, C.K. Wong, "The sequential destruction method for Monte Carlo evaluation of system reliability", Research Report of IBM Thomas J. Watson Research Center, Computer Science, RC 7337 (#31441), 1978 Sep, pp 1-16. (submitted to *IEEE Trans. Reliability*).

[10] H. Kumamoto, K. Tanaka, K. Inoue, "Efficient evaluation of system reliability by Monte Carlo method", *IEEE Trans. Reliability,* vol R-26, 1977 Dec, pp 311-315.

[11] G.S. Fishman, *Principles of Discrete Event Simulation,* John Wiley & Sons, New York, 1978.

[12] J.M. Hammersley, D.C. Handscomb, *Monte Carlo Method,* London: Methuen and Co. Ltd., 1967.

[13] L.W. Schruben, B.H. Margolin, "Psudorandom Number Assignment in Statistically Designed Simulation and Distribution Sampling Experiments", *J. American Statistical Association,* vol 73, 1978 Sep, pp 504-520.

*Available from National Technical Information Service; Springfield, VA 22161 USA.

AUTHORS

Hiromitsu Kumamoto; Dept. of Precision Mech.; Faculty of Engr.; Kyoto University; Kyoto 606 JAPAN.

H. Kumamoto: For biography, see vol R-26, 1977 Dec, p 315.

Kazuo Tanaka; Dept. of Precision Mech.; Faculty of Engr.; Kyoto University; Kyoto 606 JAPAN.

K. Tanaka: For biography, see vol R-26, 1977 Dec, p 315.

Koichi Inoue; Dept. of Precision Mechanics; Faculty of Engineering; Kyoto University; Kyoto 606 JAPAN.

K. Inoue: For biography, see vol R-23, 1974 Apr, p 33.

Ernest J. Henley; Dept. of Chem. Engr.; College of Engr.; University of Houston; Houston, Texas 77004 USA.

E. Henley: For biography, see vol R-23, 1974 Apr, p 33.

Manuscript TR79-38 received 1979 March 6; revised 1979 September 20.

★★★