Technical University of Denmark

Bounds for the probability of a union - a fault tree application

Forskningscenter Risø, Roskilde

Publication date:
1982

Document Version
Publisher's PDF, also known as Version of record

Risø-M-2340

BOUNDS FOR THE PROBABILITY OF A UNION - A FAULT TREE APPLICATION

Ole Platz

Abstract. A survey is given of methods for calculating upper and lower bounds of degree two for the probability of a union. The methods are shown to be applicable to fault tree probability calculations where they provide better bounds than the frequently used Bonferroni inequalities of degree two.

INIS Descriptors: ALGEBRA; FAULT TREE ANALYSIS; MATHEMATICS; PROBABILITY; RELIABILITY; REVIEWS

UDC 519.2:614.8

February 1982
Risø National Laboratory, DK 4000 Roskilde, Denmark.

This work was done as a part of the Nordic reactor safety project: Probabilistic Risk Analysis and Licensing (SÄK-1).

ISBN 87-550-0835-6
ISSN 0418-6435

CONTENTS

Page

1. INTRODUCTION 5

2. BOUNDS FOR THE PROBABILITY OF A UNION 6

Bounds of Degree One 6

Bounds of Degree Two 7

a) Bonferroni bounds (inclusion-exclusion bounds) 7

b) A class of upper bounds of degree two 8

c) A class of lower bounds of degree two 12

The Best Possible Bounds 13

3. A FAULT TREE APPLICATION 17

REFERENCES 21

1. INTRODUCTION

Methods for calculating the probability of a Boolean polynomial, or more precisely, for the probability of the polynomial being true, from the probabilities of its basic events are of general interest within reliability theory since fault trees for systems of binary components are just particular representations of Boolean polynomials.

For a given Boolean polynomial, the complexity of the probability calculation depends on two factors: The statistical dependency structure of its basic events, and the possibility of multiple occurrences of events in the polynomial (by this we also include the joint occurrence of an event and its negation). In the following we shall concentrate on the second factor.

If there are only few multiple occurrences of events, the probability of the Boolean polynomial may be calculated exactly by using pivotal decomposition. It is also frequently possible to calculate the probability exactly if the polynomial can be decomposed into modules, but in most cases it is necessary to use approximations which are based on the representation of the Boolean polynomial as a union of prime implicants - or minimal cut sets if the system considered is coherent.

The purpose of the present report is to give a survey of methods that have been derived for calculating upper and lower bounds of degree two for the probability of a union and to show how they may be applied to fault tree probability calculations.

2. BOUNDS FOR THE PROBABILITY OF A UNION

Let $A_1, A_2 \ldots, A_n$ be a set of arbitrary events. We consider the problem of establishing upper and lower bounds for $P(A_1 + A_2 + \ldots + A_n)$ in terms of probabilities $P(A_i)$, $P(A_iA_j)$, $P(A_iA_jA_k)$ etc. of subsets of the events. Bounds which are expressed in terms of $P(A_i)$ exclusively are called bounds of degree 1. Bounds expressed in terms of $P(A_i)$ and $P(A_iA_j)$ are called bounds of degree 2 etc. We first consider

Bounds of Degree One

These bounds, which were first derived by Boole |1|, are

$$\max[P(A_1),P(A_2), \ldots P(A_n)] \leq P(A_1+A_2+ \ldots +A_n) \leq \min[1,P(A_1)+P(A_2)+ \ldots +P(A_n)] \tag{1}$$

They were shown by Frechet |2| to be the best possible bounds if all that is known about the events $A_1, A_2 \ldots A_n$ is that their probabilities are $P(A_1), P(A_2), \ldots, P(A_n)$. Knowledge of (1) also allows us to calculate bounds for the intersection $A_1A_2 \ldots A_n$ by noting that if Φ is a Boolean polynomial of the events $A_1, A_2 \ldots, A_n$ and if L and U denote the lower and upper bound respectively for the probability of the negated event, $\bar{\Phi}$, that is if

$$L \leq P(\bar{\Phi}) \leq U \tag{2}$$

then,

$$1 - U \leq P(\Phi) \leq 1 - L \tag{3}$$

If we take $\bar{\Phi} = \bar{A}_1 + \bar{A}_2 + \ldots + \bar{A}_n$, eqs. (1) and (3) give

$$\max[0,P(A_1)+P(A_2)+ \ldots +P(A_n)-(n-1)] \leq P(A_1A_2 \ldots A_n) \leq \min[P(A_1), P(A_2), \ldots P(A_n)] \tag{4}$$

Esary and Proschan [3] showed that the upper bound in (1) could be strengthened if - in addition to the probabilities $P(A_1)$, ... $P(A_n)$ - the events A_1, A_2, ... A_n were known to be associated. Association of the events implies that

$$P(A_1 A_2 \dots A_n) \geq \prod_{i=1}^{n} P(A_i) \tag{5}$$

and since association of events implies association of the negated events, we also have

$$P(\bar{A}_1 \cdot \bar{A}_2 \dots \bar{A}_n) \geq \prod_{i=1}^{n} P(\bar{A}_i) \tag{6}$$

By using (6) and (3) we obtain the Esary Proschan bound

$$P(A_1 + A_2 + \dots + A_n) \leq 1 - \prod_{i=1}^{n} (1 - P(A_i)) \tag{7}$$

Bounds of Degree Two

a) Bonferroni bounds (inclusion-exclusion bounds)

Define the sequence S_k, k=1, 2, ..., n, by

$$S_k = \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} P(A_{i_1} A_{i_2} \dots A_{i_k}) \tag{8}$$

The following identity which is often called Poincaré's identity is known to hold:

$$P(A_1 + A_2 + \dots + A_n) = \sum_{k=1}^{n} (-1)^{k-1} S_k \tag{9}$$

A sequence of inequalities (often attributed to Bonferroni) are connected with (9):

$$P(A_1 + A_2 + \dots + A_n) \leq \sum_{k=1}^{2m-1} (-1)^{k-1} S_k \quad \text{for } m=1, 2, \dots \tag{10}$$

and

$$P(A_1+A_2+\ldots+A_n) \gtrless \sum_{k=1}^{2m} (-1)^{k-1} S_k \text{ for } m=1,\ 2,\ \ldots \tag{11}$$

The derivation of the bounds are given in several textbooks, see e.g. |4|.

The Bonferroni bounds of degree 1 and 2 are perhaps the most frequently applied bounds within reliability theory. Written out explicitly they are

$$\sum_{i=1}^{n} P(A_i) - \sum_{1\leq i<j\leq n}^{n} P(A_iA_j) \leq P(A_1+A_2+\ldots+A_n) \leq \sum_{i=1}^{n} P(A_i) \tag{12}$$

If the events A_i are the prime implicants for a system represented by a fault tree, the Bonferroni bounds of degre 2 require the probabilities of $\binom{n}{2}$ intersections A_iA_j. The bounds of degree 3 require the probabilities of $\binom{n}{3}$ intersections etc. For a fault tree with a large number of prime implicants the bounds are usually restricted to those given in (12).

The Bonferroni bounds of degree 2 may be improved upon considerably by considering a class of recently derived bounds of degree 2:

<u>b) A class of upper bounds of degree 2</u>

The derivation of the bounds is based on the following decomposition:

$$\begin{aligned} A_1+A_2+\ \ldots\ +A_n &= A_1(\Omega+A_2+\ldots+A_n)+\bar{A}_1(\emptyset+A_2+\ldots+A_n) \\ &= A_1+\bar{A}_1(A_2+\ \ldots\ +A_n) \end{aligned} \tag{13}$$

where Ω denotes the full set and $\emptyset$ the empty set.

By expanding the r.h.s. with respect to A_2 and continuing in this way, the following expression is obtained:

$$A_1+A_2+\ \ldots\ +A_n = A_1 + \sum_{i=2}^{n} (\bar{A}_1\ldots\bar{A}_{i-1}A_i) \tag{14}$$

since the terms on the r.h.s. are mutually exclusive, the probability of the union is given by

$$P(A_1+A_2+\ldots+A_n) = P(A_1)+\sum_{i=2}^{n} P(\bar{A}_1\ldots\bar{A}_{i-1}A_i) \tag{15}$$

For term No. i in the summation we have

$$P(\bar{A}_1 \ldots \bar{A}_{i-1}A_i) \leq P(\bar{A}_{i'},A_i) = P(A_i)-P(A_{i'},A_i) \tag{16}$$

where i' denotes an arbitrary choice of subscript in {1, ..., i-1}.

Insertion of (16) in (15) gives

$$P(A_1+A_2+\ldots+A_n) \leq \sum_{i=1}^{n} P(A_i) - \sum_{i=2}^{n} P(A_{i'},A_i) \tag{17}$$

The r.h.s. of (17) is thus an upper bound of $P(A_1+A_2 \ldots + A_n)$ of degree 2 which is better than the corresponding Bonferroni bound (12). The value of the bound in (17) depends on both the arbitrary choices i' and the way the events A_i are indexed.

Kounias |5| first obtained a bound of this type. His derivation is equivalent to choosing the index i' equal to 1 for all i in (9). Equation (17) then becomes

$$P(A_1+A_2+\ldots+A_n) \leq \sum_{i=1}^{n} P(A_i) - \sum_{i=2}^{n} P(A_1A_i) \tag{18}$$

Instead of starting with expansion around A_1 in (13), an arbitrary event A_k could have been selected. Instead of (18), this would give

$$P(A_1+A_2+\ldots+A_n) \leq \sum_{i=1}^{n} P(A_i) - \sum_{\substack{i=1\\ i\neq k}}^{n} P(A_kA_i) \tag{19}$$

By selecting the event k which gives the lowest value of the bound, Kounias obtained

$$P(A_1+A_2+\ldots+A_n) \leq \sum_{i=1}^{n} P(A_i) - \max_{k} \sum_{\substack{i=1\\ i\neq k}}^{n} P(A_k A_i) \qquad (20)$$

Another bound corresponding to a special selection of values i' in (17) was derived by Vanmarcke |6| and Ditlevsen |7| who applied it to problems within structural reliability. Their derivation is equivalent to choosing for each term in the second sum in the r.h.s. of (15) the value of i' giving the maximum value of $P(A_i,A_i)$. The bound obtained by this selection is

$$P(A_1+A_2+\ldots+A_n) \leq \sum_{i=1}^{n} P(A_i) - \sum_{i=2}^{n} \max_{k<i} P(A_k A_i) \qquad (21)$$

In the following the r.h.s. of (20) will be denoted the MS--bound (maximal sum), and the r.h.s. of (21) the SM-bound (sum of maxima). Both bounds are easily calculated but neither of them is necessarily the best (i.e. minimum) among the bounds represented by the general expression (17).

Hunter |8| derived a graph theoretical method for calculating the minimum of the bounds represented by the r.h.s. of (17). Assume that each of the events A_i correspond to a vertex of a graph and let the intersections A_iA_j represent edges (denoted (i,j)) joining two of the vertices. A spanning tree of the vertices A_i, i=1, ..., n, is a connected graph without cycles with n-1 edges such that at least one edge is incident on each of the k vertices. Hunter proved that:

For some assignment of subscripts and some arbitrary choices, i', a set of n-1 intersections may be used in the second term of (17) if and only if it forms a spanning tree of the vertices $\{A_i\}$.

With this result (17) is equivalent to

$$P(A_1+A_2+\ldots+A_n) \leq \sum_{i=1}^{n} P(A_i) - \sum_{\tau} P(A_i A_j) \qquad (22)$$

where $\tau = \{(i,j)\}$ is an arbitrary choice of one of the n^{n-2} possible spanning trees among the n vertices (A_i). If the set of all spanning trees among the n vertices is denoted T, the lowest upper bound for $P(A_1+A_2+ \ldots +A_n)$ is

$$P(A_1+A_2+\ldots+A_n) \leq \sum_{i=1}^{n} P(A_i) - \max_{\tau \in T} \sum_{\tau} P(A_i A_j) \qquad (23)$$

The bound in (23) is better than both the MS-bound (20) and the SM-bound (21) since these are the minima of n and (n-1)! respectively, of the total of n^{n-2} spanning trees in T.

As a simple illustration consider the case of n=4. The n^{n-2} = 16 spanning trees of the graph with vertices A_1 to A_4 are shown in figure 1. The spanning trees covered by (20) and (21) with this indexing of events are marked MS and SM respectively.

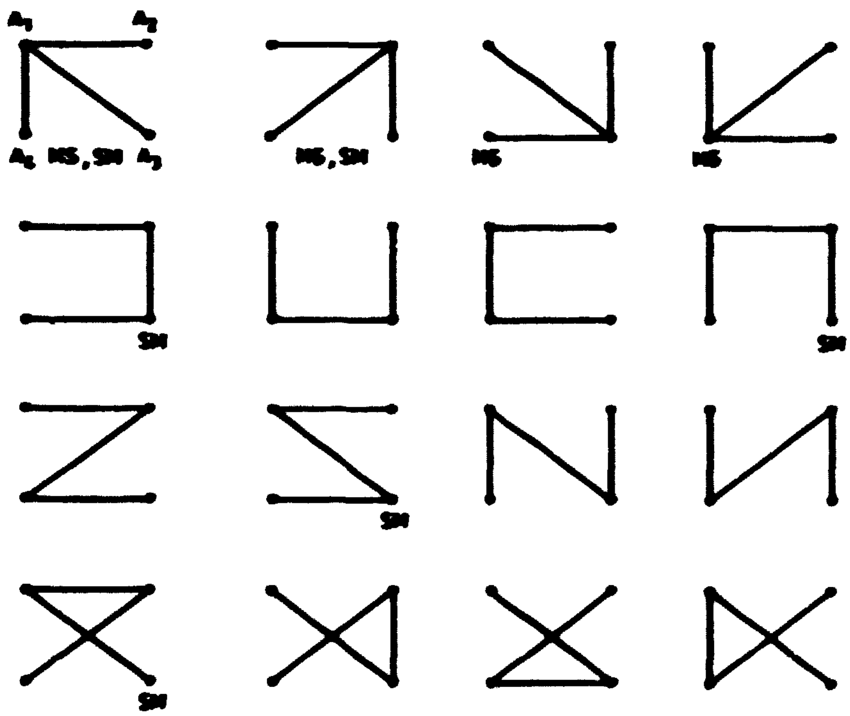

Figure 1. The 16 spanning trees among the vertices $A_1 \ldots A_4$. The spanning trees covered by (20) and (21) are marked MS and SM respectively.

Calculation of the tree τ' that maximizes the second term of the r.h.s. of (16) can be done by considering $P(A_iA_j)$ to be the weight of the edge (i, j) and using Kruskal's |9| algorithm for finding the maximal spanning tree. Let U denote the set of unconnected vertices and C the set of connected vertices and let initially $C=\emptyset$ and $U=\{A_i\}$. Transfer an arbitrary vertex A_i from U to C. The algorithm as stated by Hunter is:

1) Find the largest $P(A_iA_j)$ such that $A_i \epsilon C$ and $A_j \epsilon U$. Let the branch corresponding to this intersection be (i', j').

2) Include (i', j') in τ' and remove A_j' from U and place it in C. If $U \neq \emptyset$ go to 1), otherwise, τ' is complete.

The procedure yields τ' in n-1 steps.

<u>c) A class of lower bounds of degree two</u>

Let J_r denote a subset of $\{1, 2, \ldots, n\}$ with r elements i_1, $i_2, \ldots, i_r$. Since

$$A_{i_1}+A_{i_2}+\ldots+A_{i_r} \subseteq A_1+A_2+\ldots+A_n \tag{24}$$

then

$$P(A_{i_1}+A_{i_2}+\ldots+A_{i_r}) \leq P(A_1+A_2+\ldots+A_n) \tag{25}$$

By using the Bonferroni inequality of degree two, the left part of eq. (12), to $P(A_{i_1}+A_{i_2}+\ldots+A_{i_r})$ gives

$$P(A_1+A_2+\ldots+A_n) \geq \sum_{i\epsilon J_r} P(A_i) - \sum_{\substack{i<j \\ i,j\epsilon J_r}} P(A_iA_j) \tag{26}$$

By taking the maximum over all sets J_r, we obtain the lower bound derived by Kounias |5|:

$$P(A_1+A_2+\ldots+A_n) \geq \max_{J_r} \Big(\sum_{i\epsilon J_r} P(A_i) - \sum_{\substack{i<j \\ i,j\epsilon J_r}} P(A_iA_j) \Big) \tag{27}$$

The best possible bounds

Boole |1| formulated the general problem of finding the best possible bounds for the probability of a logical polynomial given the probabilities of other polynomials of the same basic events. Hailperin |10| showed that the problem could be formulated and solved as a primal or dual linear programming problem. Kounias and Marin |11| found the basic feasible solutions for the dual linear programming problem for the probability of a union of events A_1, A_2, ..., A_n for n=2, 3 and 4 given that $P(A_i)$ and $P(A_iA_j)$ were known. They also found some of the feasible solutions for $n \geq 5$, but these bounds are not as easily calculated as the bounds described in the previous sections.

Example

As an example, consider a set of 10 events with probabilities $P(A_i)$ and $P(A_iA_j)$ given in figure 2. The probabilities $P(A_i)$ are shown in the diagonal. Due to the symmetry of the matrix, the elements below the diagonal are not shown.

i \ j	1	2	3	4	5	6	7	8	9	10
1	.180	.005	.022	.027	.036	.054	.027	.045	.045	.007
2		.030	.012	.015	.006	.009	.009	.005	.005	.001
3			.120	.060	.024	.036	.036	.018	.018	.012
4				.150	.030	.045	.045	.023	.023	.012
5					.200	.060	.018	.060	.030	.008
6						.300	.054	.090	.090	.024
7							.090	.045	.045	.007
8								.150	.075	.012
9									.150	.012
10										.040

Fig. 2. Matrix of probabilities $P(A_iA_j)$.

The upper bound of degree one is trivially 1 since

$$\sum_{i=1}^{10} P(A_i) = 1.410 \tag{28}$$

The MS-bound, eq. (20) is 1.410 minus the sum of the off--diagonal elements in the 6th column and the 6th row in figure 2. Its value is

$$\text{MS-bound} : 1.410 - 0.462 = 0.948 \tag{29}$$

The SM-bound, eq. (21), is obtained by finding the largest off-diagonal element in each column and subtracting their sum from 1.410.

$$\text{SM-bound} = 1.410 - 0.441 = 0.969 \tag{30}$$

The algorithm for finding the spanning tree (ST) bound may be implemented directly on the matrix of probabilities $P(A_iA_j)$. It is convenient to start with the matrix in a slightly different form compared to figure 2: Delete the diagonal elements and fill in the elements below the diagonal. Then proceed as follows,

Select an arbitrary row, encircle its row index and delete the column with the same index.

Find the largest element in the row with the encircled index. Encircle it and delete the rest of the elements in that column. Encircle the row index which is equal to the index of the column just deleted. The situation is now as shown in figure 3.

i \ j	1	2	3	4	5	6	7	8	9	10
(1)		.005	.022	.027	.036	(.054)	.027	.045	.045	.007
2	.005		.012	.015	.006	.009	.009	.005	.005	.001
3	.022	.012		.060	.024	.036	.036	.018	.018	.012
4	.027	.015	.060		.030	.045	.045	.023	.023	.012
5	.036	.006	.024	.030		.060	.018	.060	.030	.008
(6)	.054	.009	.036	.045	.060		.054	.090	.090	.024
7	.027	.009	.036	.045	.018	.054		.045	.045	.007
8	.045	.005	.018	.023	.060	.090	.045		.075	.012
9	.045	.005	.018	.023	.030	.090	.045	.075		.012
10	.007	.001	.012	.012	.008	.024	.007	.012	.012	

Fig. 3.

To proceed, select the greatest element among the elements in the rows with encircled indexes. Encircle the largest element, delete the rest of the elements in that column, and encircle the row index which is equal to the index of the column just deleted. Proceed in this way until all columns are deleted (figure 4).

In the course of the process, the encircled indexes correspond to the vertices that are placed in the set C in the spanning tree algorithm. Deletion of a column corresponds to removal of a vertex from U to C as soon as an edge has been connected to that vertex.

i \ j	1	2	3	4	5	6	7	8	9	10
1		.005	.022	.027	.036	**.054**	.027	.045	.045	.007
2	.005		.012	.015	.006	.009	.009	.005	.005	.001
3	.022	.012		.060	.024	.036	.036	.018	.018	.012
4	.027	**.015**	**.060**		.030	.045	.045	.023	.023	.012
5	.036	.006	.024	.030		.060	.018	.060	.030	.008
6	.054	.009	.036	**.045**	**.060**		**.054**	**.090**	**.090**	**.024**
7	.027	.009	.036	.045	.018	.054		.045	.045	.007
8	.045	.005	.018	.023	.060	.090	.045		.075	.012
9	.045	.005	.018	.023	.030	.090	.045	.075		.012
10	.007	.001	.012	.012	.008	.024	.007	.012	.012	

Fig. 4.

Finally, the spanning tree bound is obtained by subtracting from 1.410 the sum of the encircled matrix elements

$$\text{ST-bound}: 1.410 - 0.492 = 0.918 \tag{31}$$

For the calculation of the lower bound from equation (26), the set $J_4 = \{1, 4, 5, 6\}$ gives

$$\sum_{i \in J_4} P(A_i) - \sum_{\substack{i<j \\ i,j \in J_4}} P(A_i A_j) = 0.578 \tag{32}$$

which is much better than the lower Bonferroni bound of degree two:

$$\sum_{i \in J_{10}} P(A_i) - \sum_{\substack{i<j \\ i,j \in J_{10}}} P(A_i A_j) = 0.068 \tag{33}$$

3. A FAULT TREE APPLICATION

The methods described in the previous section may be applied to provide bounds for the top event probability of a fault tree if the sets A_i are taken to be the prime implicants of the systems failure polynomial. The bounds are easily calculated since they do not require probabilities beyond those already calculated for the Bonferroni bounds.

As an illustration consider the fault tree in figure 5. The prime implicants - which are minimal cut sets in this case since there are no negated events - are

$$A_1 = 1 . 11$$
$$A_2 = 2 . 6$$
$$A_3 = 2 . 8$$
$$A_4 = 2 . 12$$
$$A_5 = 3 . 4$$
$$A_6 = 5 . 10$$
$$A_7 = 2 . 5 . 11$$
$$A_8 = 3 . 5 . 11$$
$$A_9 = 5 . 7 . 11$$
$$A_{10} = 5 . 8 . 9 . 12$$

The basic events are assumed to be statistically independent. If the probability of each of them is 0.1, the following probabilities for $P(A_i)$ and $P(A_iA_j)$ are obtained:

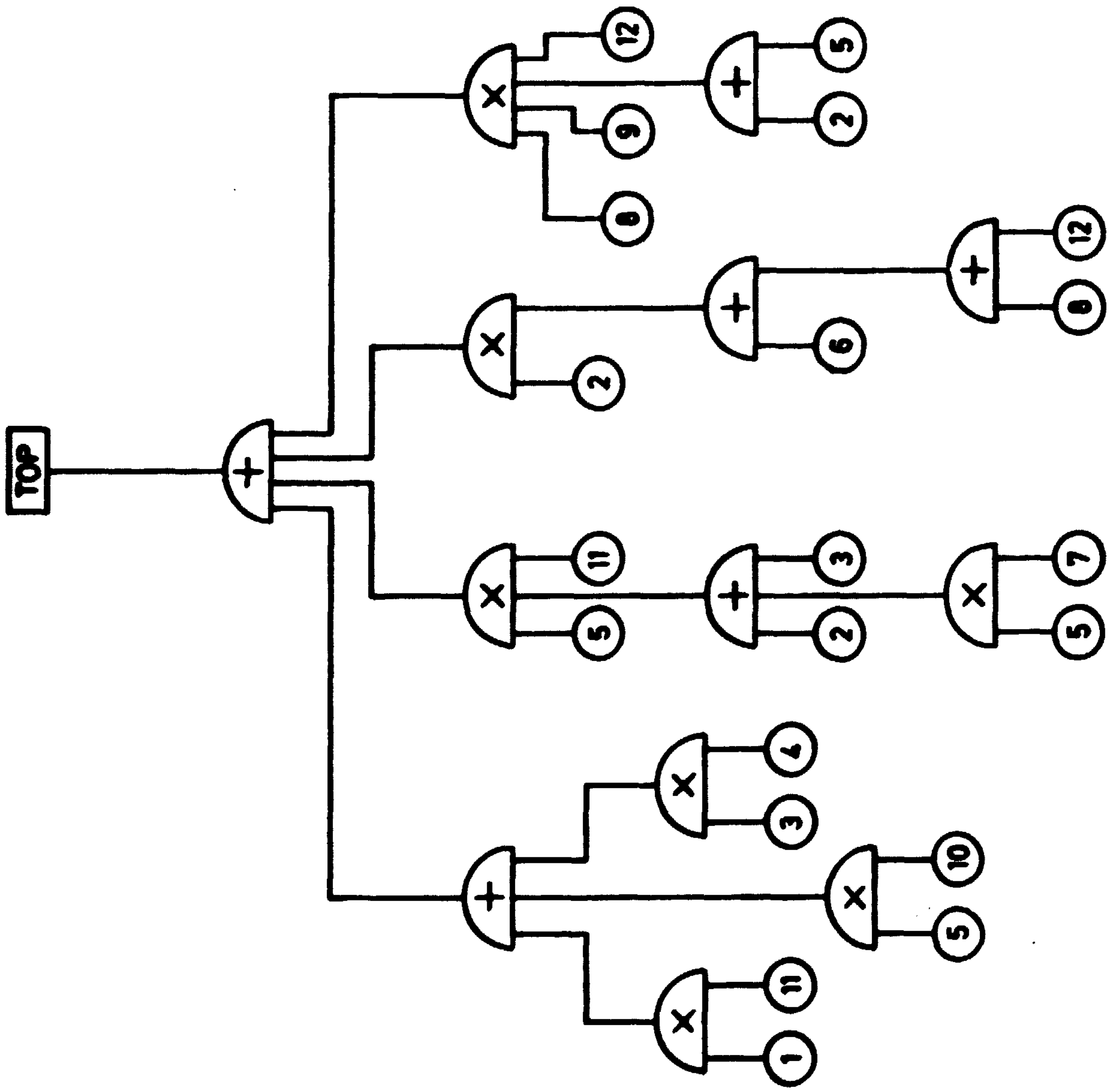

Fig. 5.

i \ j	1	2	3	4	5	6	7	8	9	10
1	10^{-2}	10^{-4}	10^{-4}	10^{-4}	10^{-4}	10^{-4}	10^{-4}	10^{-4}	10^{-4}	10^{-6}
2		10^{-2}	10^{-3}	10^{-3}	10^{-4}	10^{-4}	10^{-4}	10^{-5}	10^{-5}	10^{-6}
3			10^{-2}	10^{-3}	10^{-4}	10^{-4}	10^{-4}	10^{-5}	10^{-5}	10^{-5}
4				10^{-2}	10^{-4}	10^{-4}	10^{-4}	10^{-5}	10^{-5}	10^{-5}
5					10^{-2}	10^{-4}	10^{-5}	10^{-4}	10^{-5}	10^{-6}
6						10^{-2}	10^{-4}	10^{-4}	10^{-4}	10^{-5}
7							10^{-3}	10^{-4}	10^{-4}	10^{-6}
8								10^{-3}	10^{-4}	10^{-6}
9									10^{-3}	10^{-6}
10										10^{-4}

Fig. 6. Probabilities $P(A_i A_j)$ for the fault tree example.

The upper bounds for the top event probability are shown in table 1.

Table 1. Upper probability bounds (multiplied by 10^2), for the fault tree example. The abbreviations are; BF: Bonferroni (12), MS: Maximal sum (20), SM: Sum of maxima (21), ST: Spanning tree (23), EP: Esary and Proschan's bound (7).

BF	MS	SM	ST	EP
6.31	6.07	6.05	6.05	6.14

In this case, the best lower bound among the bounds given in (27) is the Bonferroni bound (12). Its value is $5.75 \cdot 10^{-2}$.

For comparison, the exact probability, which may be obtained by pivotal decomposition of the Boolean polynomial, is $5.80 \cdot 10^{-2}$.

REFERENCES

1. G. Boole, Laws of Thought, reprint of 1854 ed. Dover, New York.
2. M. Fréchet, Généralizations du théoréme des probabilitiés totales, Fund. Math. 25 (1935) 379-387.
3. J. D. Esary and F. Proschan, A reliability bound for systems of maintained, interdependent components. J. Amer. Statist. Assoc. 65 (1970) 329-338.
4. W. Feller, An introduction to probability theory and its applications. Volume I, 3rd ed. John Wiley and Sons, New York (1968).
5. E. Kounias, Bounds for the probability of a union, with applications. Ann. Math. Statist. 39 (1968) 2154-2158.
6. E. H. Vanmarcke, Matrix formulation of reliability analysis and reliability-based design. Computers & Structures 3 (1971) 757-770.
7. O. Ditlevsen, Narrow reliability bounds for structural systems. J. of Struct. Mech. 7 No. 4 (1979).
8. D. Hunter, An upper bound for the probability of a union. J. Appl. Prob. 13 (1976) 597-603.
9. J. B. Kruskal, On the shortest spanning subtree of a graph and the travelling salesman problem. Proc. Amer. Math. Soc. 7 (1956) 48-50.
10. T. Hailperin, Best possible inequalities for the probability of a logical function of events. Am. Math. Month. 72 (1965) 343-359.
11. S. Kounias and J. Marin, Best linear Bonferroni bounds. SIAM J. Appl. Math. 30 (1976) 307-323.

Title and author(s)

BOUNDS FOR THE PROBABILITY OF A UNION -
A FAULT TREE APPLICATION

Ole Platz

Date March 1982

Department or group

Electronics

Group's own registration number(s)

R-5-82

SÅK-1-D(82)-1

OP/AME

pages + tables + illustrations

Abstract

A survey is given of methods for calculating upper and lower bounds of degree two for the probability of a union. The methods are shown to be applicable to fault tree probability calculations where they provide better bounds than the frequently used Bonferroni inequalities of degree two.

Copies to

Available on request from Risø Library, Risø National Laboratory (Risø Bibliotek), Forsøgsanlæg Risø), DK-4000 Roskilde, Denmark
Telephone: (03) 37 12 12, ext. 2262. Telex: 43116