Fault Dictionary Size Reduction through Test Response Superposition

Baris Arslan and Alex Orailoglu
Computer Science and Engineering Department
University of California, San Diego
La Jolla, CA 92093
{barslan, alex}@cs.ucsd.edu

Abstract

The exceedingly large size of fault dictionaries constitutes a fundamental obstacle to their usage. We outline a new method to reduce significantly the size of fault dictionaries. The proposed method partitions the test set and a combined signature is stored for each partition. The new approach aims to provide high diagnostic resolution with a small number of combined signatures. The experimental results show a considerable decrease in the storage requirement of fault dictionaries.

1. Introduction

Improvement of the VLSI IC production quality relies fundamentally on the reduction of the number of defective circuits during the circuit manufacturing process. Production quality improvement depends on fault diagnosis for determining the cause and location of the failures. Diagnosis uses the response of a circuit under test to find the physical location of failure. Knowledge regarding fault location is essential to the improvement of the design and manufacturing process.

Diagnosis approaches can be partitioned into two fundamental alternatives: diagnosis using pre-computed fault dictionaries and dynamic diagnosis [1], [13]. Diagnosis requires internal knowledge of a circuit under test. The fault dictionary stores the responses of a circuit to a test set in the presence of faults under a fault model. Diagnosis using a fault dictionary compares the observed response of the circuit under test with the pre-computed responses in the fault dictionary. Dynamic diagnosis analyzes the response of the CUT to determine the actual cause of error. The typical objections to the usage of fault dictionaries consist of their exceedingly large size and the high computational time required for dictionary creation. However, creation of a fault dictionary for a particular circuit is performed only once. Pomeranz and Reddy [11] show that only a small number of diagnostic runs on circuit copies suffices to compensate for the additional nonrecurring computational cost of fault dictionary based diagnosis over dynamic diagnosis. The main objection to the usage of fault dictionaries nonetheless remains their large sizes. Increasing circuit sizes push the storage requirement of fault dictionaries to unacceptable levels.

The diagnosis process aims at distinguishing fault pairs. If some faults exhibit the same response to a test set, they form an equivalence class, limiting the diagnostic resolution of the test set [1]. In order to distinguish all possible fault pairs, the diagnosis process has to partition all faults into distinct singleton sets. If each test pattern in a test set evenly divides the current fault partitions, this goal can be achieved with a test set that is optimal in terms of test set size. Since the fault dictionary stores the response of the circuit to the test set in the presence of each fault, decreasing the number of test vectors in the test set leads to smaller fault dictionary sizes.

We propose a method to decrease the size of a fault dictionary. In the proposed method, no modification to the test pattern generator is undertaken, retaining the viability of utilizing the proposed method in current industrial test generation practices. We use the responses of the CUT to test patterns to form signatures that partition the faults as evenly as possible. We subsequently partition the test set to small distinct test sets, wherein one combined response for each test set partition is stored and high resolution is achieved with the minimum number of partitions.

In this paper, we focus on the aforementioned problem of fault dictionaries, propose a method to reduce the large datasets and show the results of the presented solution through experiments. Section 2 outlines previously proposed methods for dictionary size reduction. Section 3 discusses a set of observations that lead to the method proposed in this paper. Section 4 presents our dictionary compaction method; the suggested method is experimentally evaluated in Section 5. Conclusions are drawn in Section 6.

2. Previous Work

Diagnosis using fault dictionaries consists of applying test vectors to the CUT and comparing the responses observed at primary outputs to the stored responses in the fault dictionary in order to identify faults that map to the observed response.

A full fault dictionary stores for each test vector the responses at all outputs of the circuit in the presence of each modeled fault. While a full fault dictionary captures all the available information, its size is exceedingly large; in the presence of R primary outputs, T test vectors and F faults, the size of a full fault dictionary is O($F * T * R$). In order to reduce the space needed to store this exceedingly large data set, many methods have been proposed [11], [8], [2], [3], [6]. An extensive summary can be found in [12].

A pass-fail dictionary stores for each test vector only pass or fail information in the presence of each modeled fault. Although diagnostic resolution of a pass-fail dictionary lags that of a full fault dictionary, for circuits with a large number of primary outputs, its size, O($F * T$), is sharply reduced compared to that of the corresponding full dictionary.

A straightforward compaction technique, industrially popular, is the detection dictionary that stores the outputs of failed test vectors only for each modeled fault, thus reducing size but retaining essentially the same information as a full fault dictionary. Similarly a vector dictionary typically compacts the pass-fail dictionary by storing only the index of the failing test vectors for each fault. Both of these techniques are reviewed in [6].

Pomeranz and Reddy [11] propose the *Compact* dictionary. The *Compact* dictionary, essentially analogous to a pass-fail dictionary, exhibits comperatively increased resolution by allowing pass/fail responses to be computed on the subsets of outputs for some test vectors. The aim is to partition the outputs so that the minimal number of groups is used to achieve the resolution of a full fault dictionary.

Lavo and Larrabee [8] present a dictionary structure, wherein in addition to storing a pass-fail signature, they include *output-compacted* signatures for each fault in the dictionary. In the output-compacted signature, each bit indicates whether a particular output ever fails under the test set in the presence of the modeled faults, adding an additional dimension of useful failing output information to the pass-fail dictionary.

Chess and Larrabee [6] present a dictionary organization based on error sets. The error set is a set of primary outputs that manifest errors. Since distinct test vectors may produce the same error set in the presence of a fault, the proposed dictionary organization stores *(vector set, error set)* pairs for each fault, removing inherent redundancy in the dictionary.

Pass-fail and vector dictionaries impose the smallest storage requirements but their diagnostic resolution is typically worse. For T test vectors, $\lceil log_2 T \rceil$ bits are needed to represent test vectors in the vector dictionary. Consequently, if the number of test vectors that detects fault f_i is n_i, $n_i * \lceil log_2 T \rceil$ bits are required for fault f_i. Pomeranz and Reddy [11] show for the circuits in ISCAS85 benchmark [4] that the size of vector dictionaries exceeds that of pass-fail dictionaries.

3. Motivation

The large size and high computational time for dictionary creation constitute the fundamental problems of fault dictionaries. As mentioned in Section 1, Pomeranz and Reddy [11] show that the superiority of diagnosis using a fault dictionary over dynamic diagnosis in terms of computational time becomes evident after only a small number of diagnosis runs. Consequently, the main focus in this paper is the remaining issue of the exceedingly large size of the fault dictionaries.

While a measure of fault resolution is typically lost in using a pass-fail dictionary, the drastic decrease in storage cost makes its cost-benefit tradeoff highly appealing. Nonetheless, the storage cost implications of current industrial ICs and associated diagnosis problems preclude the applicability even of pass-fail dictionaries for large circuits. It is essential that approaches for further drastic reductions in dictionary size be exploited, possibly even at the expense of slight fault resolution degradation, utilizing the same advantageous tradeoff that led to the adoption of pass-fail dictionaries to satisfy the stringent constraints imposed by the ICs of prior generations.

The diagnosis process aims at distinguishing all possible pairs of faults. In order to achieve this goal, the diagnosis process has to partition all faults to distinct sets. Application of each test vector in the test set partitions the set of indistinguishable faults to smaller equivalence sets. Formally, the faults that continue to be indistinguishable after the application of the full test vector suite are equivalent under the currently applied test set [1]. In order to provide full resolution for the modeled faults, each fault should be in a distinct equivalence class at the end of the application of the test set.

Assuming F modeled faults in the system, at best $\lceil log_2 F \rceil$ test vectors suffice to fully distinguish all faults in the CUT. However, in practice the number of ATPG-generated test vectors to be used for diagnosis is quite a bit higher than the minimal theoretical bounds for full distinguishability of all faults as can be seen experimentally by comparing the second and fourth columns in Table 1 for a number of ISCAS85 benchmark circuits. At best, each test vector partitions the indistinguishable fault sets evenly.

Circuit	$\|T\|$	$\|F\|$	$\lceil log_2 F \rceil$	P/F Dict.	XORed
c432	51	524	10	0.117	0.453
c499	53	758	10	0.352	0.453
c880	59	942	10	0.272	0.486
c1355	84	1574	11	0.194	0.472
c1908	115	1879	11	0.242	0.485
c2670	110	2747	12	0.196	0.460
c3540	158	3428	12	0.131	0.450
c5315	121	5350	13	0.154	0.475
c6288	34	7744	13	0.351	0.475
c7522	210	7550	13	0.173	0.472
Average				0.219	0.469

Table 1. The fraction of detected faults per test vector

Therefore, in order to reach the maximum diagnostic resolution goal with the minimum number of test vectors, the number of detected faults (1's in the columns of the pass-fail dictionary) has to approach the number of undetected faults (0's in the columns of the pass-fail dictionary) for each test vector. However, the fraction of faults detected by a test vector is typically quite small. To shed a bit of empirical data on this question, each test vector, for example, detects on the average 21.9% of the modeled faults for the circuits in the ISCAS85 benchmark set as can be seen in the third column of Table 1.

Each new test vector created by ATPG targets a fault that has not hitherto been detected by previous test vectors. Therefore, although there will be overlaps for easy-to-detect faults among the responses of test vectors, the overlaps in the responses are expected to be fewer for hard-to-detect faults. Since the detected faults for each test pattern are quite a bit fewer than undetected faults and since test vectors target distinct faults, XORing the pass-fail responses of various test vectors promises to increase the number of 1's in the new signature, thus approaching an even distribution of 1's and 0's. While the best distribution typically is to be achieved when partitions need not obey strict test vector sequence constraints, even a straightforward partitioning approach, respecting the actual order of the test vector sequence, increases the fraction of 1's to 46.9% for the circuits in ISCAS85 on average as can be observed in the last column of Table 1.

If XORing the columns of a pass-fail dictionary provides a better distribution than individual columns of the pass-fail dictionary, the test set can be partitioned to smaller test sets and the responses of the test vectors within the new smaller test sets can be XORed to produce a single combined response for each test vector partition. Since the new combined responses provide an improved distribution compared to individual test vector responses, it is possible to provide high resolution with fewer combined responses than individual test vector responses.

If the test set is partitioned to smaller test sets and a combined response is needed for each test vector partition, only one column for each test vector partition will suffice in the new dictionary. Therefore, the ultimate goal is the identification of an optimal partitioning of test vectors that enable high resolution, based on the single combined response for each partition.

4. Algorithm

In the first phase of the algorithm, to attain the goal of an even distribution, the XORed pass-fail dictionary is created. The new dictionary is very similar to the pass-fail dictionary, wherein each row represents a modeled fault, f_i, and each column represents a test vector, t_i. In the XORed pass-fail dictionary, each row represents a modeled fault as in the pass-fail dictionary but the *ith* column of the XORed pass-fail dictionary is the XOR of the first *i* columns of the pass-fail dictionary of the same circuit. Formally, given a test vector sequence $T =< t_1, t_2, \ldots, t_n >$, let the *ith* prefix of T be $T_i =< t_1, t_2, \ldots, t_i >$. The *ith* column of the pass-fail dictionary provides information regarding whether the test vector, t_i, detects a fault. The *ith* column of the XORed pass-fail dictionary, on the other hand, provides the XOR of the columns in the pass-fail dictionary for test vectors in the *ith* prefix of T, T_i. For a given pass-fail dictionary of a circuit with four modeled faults, $F = \{f_1, f_2, f_3, f_4\}$ and four test vectors, $T = \{t_1, t_2, t_3, t_4\}$ shown in Figure 1, the corresponding XORed pass-fail dictionary is shown in Figure 2.

The next step consists of the creation of *the distinguishability table*, wherein each row represents a fault pair, (f_i, f_j) and the *kth* column of row (f_i, f_j) indicates whether f_i and f_j are distinguishable through an XOR of the columns in the pass-fail dictionary for the test vectors in T_k (O(t_1)$\oplus$O(t_2)$\oplus \cdots \oplus$O(t_k)). In other words, the *kth*

	t_1	t_2	t_3	t_4
f_1	1	0	0	0
f_2	0	1	0	1
f_3	0	0	0	1
f_4	0	0	1	1

Figure 1. Pass-Fail Dictionary

	T_1	T_2	T_3	T_4
f_1	1	1	1	1
f_2	0	1	1	0
f_3	0	0	0	1
f_4	0	0	1	0

Figure 2. XORed Pass-Fail Dictionary

	T_1	T_2	T_3	T_4
(f_1, f_2)	1	0	0	1
(f_1, f_3)	1	1	1	0
(f_1, f_4)	1	1	0	1
(f_2, f_3)	0	1	1	1
(f_2, f_4)	0	1	0	0
(f_3, f_4)	0	0	1	1

Figure 3. Distinguishability Table

column of row (f_i, f_j) of the distinguishability table is 1 if the entries (f_i, T_k) and (f_j, T_k) of an XORed pass-fail dictionary assume different values. Figure 3 illustrates the distinguishability table for the XORed pass-fail dictionary of Figure 2.

In the second phase, the problem of finding the optimal number of test vector partitions is mapped to a set cover problem. The set cover problem consists of a finite set of elements, F, and a set, S, of some subsets of F. The set cover problem aims at finding the minimal number of elements of S, S_i, such that the union of them covers all the elements in F.

In the process of determining the minimal number of test vector partitions, the fault pair set, wherein each fault pair is represented by a row in the distinguishability table, is mapped to the set F in the set cover problem. The fault pairs that can be distinguished by the XOR of the responses of test vectors in T_i (fault pairs with 1 in the column T_i of the distinguishability table) map to the elements of the set S_i in the set cover problem. Therefore, the minimal set of columns that covers all the rows in the distinguishability table is the optimal partitioning. If the *ith* and *jth* columns, $i < j$, are included in the minimal set with no intervening columns included, then test vectors $t_{i+1}, t_{i+2}, \cdots, t_j$ form a partition.

The NP-completeness of the set cover problem [7] necessitates the consideration of possibly sub-optimal but computationally effective algorithms; a straightforward greedy algorithm shown below is used in the experiments we conduct:

GREEDY-SET-COVER()
$U \leftarrow F$
$C \leftarrow \emptyset$
while $U \neq \emptyset$
do select $X \in S$ that maximizes $|S \cap U|$
$U \leftarrow U - X$
$C \leftarrow C \cup \{X\}$
return C

Essentially, this greedy algorithm selects the column of the distinguishability table that covers the largest number of rows and adds this column to the minimal set. Then it selects the next column that covers the largest number of rows from the remaining uncovered rows and repeats this process until all rows have been covered. For example, in Figure 3, T_1 and T_3 cover three rows and T_2 and T_4 cover four rows. Since the column that covers the largest number of rows is selected, either T_2 or T_4 is placed into the minimal set. Since no priority is set in the case of a tie, we arbitrarily select T_2 for illustration purposes. The remaining uncovered fault pairs are (f_1, f_2) and (f_3, f_4). T_1 and T_3 cover one of the pairs but T_4 covers both. Consequently, T_4 is selected as the next column and since all rows are covered thereupon, the algorithm terminates. Only two partitions $\{t_1, t_2\}$ and $\{t_3, t_4\}$ suffice to distinguish all fault pairs, translating into a 50% reduction in the size of dictionary for the example pass-fail dictionary of Figure 1, with four columns. Noteably, the same algorithm applied without creating the XORed pass-fail table necessitates at least three test vectors to provide the same level of coverage.

5. Experimental Results

The proposed method has been applied to all the circuits in ISCAS85 benchmark set. The ATALANTA test generation tool [9] and the HOPE fault simulation tool [10] have been used for the experiments. In the first set of experiments, the test patterns are generated with no utilization of any test compaction method. The method we propose has been applied on the generated test sets subsequently. Table 2 lists the associated results; the second column denotes the number of test vectors with no test compaction, the third column denotes the number of faults, the fourth column the number of partitions and the last column the number of test vectors with test compaction. The application of the method results in significant improvement as can be seen in Table 2, with a size reduction exceeding 40% on average and even more interestingly exceeding 60% for the larger

Circuit	$\|T\|$	$\|F\|$	Partitions	$\|CompactT\|$
c432	76	524	57	52
c499	70	758	51	53
c880	115	942	45	52
c1355	113	1574	81	85
c1908	187	1879	101	121
c2670	195	2747	96	104
c3540	261	3428	131	153
c5315	222	5350	86	120
c6288	55	7744	31	32
c7552	366	7550	120	207

Table 2. ISCAS85 Partition Results and Comparison to Compacted Test Vectors

Circuit	$\|T\|$	$\|F\|$	Number of Test Vector Partitions	Diagnostic Resolution using a Pass-Fail Dictionary	Diagnostic Resolution with Partitioning
c432	52	524	10	0.997971	0.993403
c499	54	758	10	0.999495	0.997257
c880	52	942	10	0.999598	0.996970
c1355	86	1574	11	0.999246	0.997274
c1908	121	1879	11	0.999418	0.997063
c2670	111	2747	12	0.997813	0.996088
c3540	154	3428	12	0.998301	0.996694
c5315	119	5350	13	0.999769	0.999083
c6288	32	7744	13	0.999806	0.999555
c7552	207	7550	13	0.999564	0.998871
Average				0.999098	0.997226

Table 3. The Diagnostic Resolution of the Proposed Method

ISCAS85 circuits. As typical industrial practice exploits aggressively compaction techniques, we have also applied our techniques on compacted vectors. The resulting highly constrained fault distinguishability relationships in the compact test set preclude the possibility of significant storage reduction, particularly in the case when hard-to-detect faults predominate. Yet interestingly, the resulting number of partitions is even smaller than the number of test vectors that would have been created, had compaction methods been applied, thus attaining simultaneously the goals of reduced storage and increased fault resolution. While a slight deterioration in test application time may be suffered in the process, its relative importance pales given the limited number of diagnosis runs per IC. To provide an example, although the number of test vectors in the compacted test set (equating the number of columns in the pass-fail dictionary) is 207 for the c7552 benchmark circuit as can be seen in the last column of Table 2, it is 120 in the proposed dictionary created using the test set without compaction; the further reduction in the dictionary size exceeds 40%.

In the next set of experiments, instead of trying to perfectly match the *diagnostic resolution*[1] of a pass-fail dictionary, the diagnostic resolution of the proposed dictionary utilizing $\lceil log_2 F \rceil$ partitions is evaluated. Table 3 lists the results; the second, third and fourth columns denote the number of test vectors, faults and test vector partitions respectively, the fifth column denotes the diagnostic resolution using a pass-fail dictionary and the last column denotes the diagnostic resolution using the proposed approach.

An additional set of experiments, shown in Table 4, is used to assess the possible diagnostic resolution loss as a result of XORing for the technique we propose. While the first four columns are analogous to those reported in Table 3, the penultimate column shows the loss in diagnostic resolution when $\lceil log_2 F \rceil$ partitions each storing the XORed response are employed, compared to the general pass-fail dictionary utilizing $|T|$ vectors. This diagnostic resolution loss can be compared to the one shown in the last column, which reports the ratio of the diagnostic resolution loss of the best $\lceil log_2 F \rceil$ unaltered responses compared to those of the pass-fail dictionary utilizing $|T|$ vectors.

When we look at the penultimate column, we see that the diagnostic resolution ratio is on the average 99.81%. The ratio exceeds 99.9% for large circuits, indicating a drastic improvement in dictionary size reduction for the larger circuits. The reduction in the number of columns in the pass-fail dictionary from $|T|$ to $\lceil log_2 F \rceil$ is achieved with negligible decrease in resolution. A close look at the experimental results of the individual circuits evinces increasingly superior results as the sizes of the circuits get larger. This improvement results from the fact that as the number of test vectors increase, the possibilities for different partitionings of the test set increase. In addition, since the number of partitions is $\lceil log_2 F \rceil$, it indicates a drastic decrease in dictionary size for larger circuits. For example, the size of the resultant dictionary is 6.2% of the pass-fail dictionary, a sixteen-fold decrease, for the largest circuit in the benchmark, c7552, while the diagnostic resolution of the new dictionary is 99.93% of the pass-fail dictionary, suffering a resolution loss of less than a thousandth.

In contrast, as can be seen from the last column of Table 4, 94.95% of diagnostic resolution of the pass-fail dictionary is achieved on average when the best, but non-XORed, $\lceil log_2 F \rceil$ responses are selected. The diagnostic resolution of the two simple pass-fail dictionaries is close, irrespective of the number of test vectors, for the c499 and c6288 benchmark circuits, reflecting the fact that the fraction of detected faults for each vector on average is already close to an even distribution, as can be observed in Table 1.

Experimental results indicate that the proposed approach

[1] We adopt the definition given in [5], which defines diagnostic resolution as the the ratio of distinguished to total fault pairs.

Circuit	$\|T\|$	$\|F\|$	Number of Test Vector Partitions	DR Ratio for XORed P/F Dictionary	DR Ratio for minimal P/F Dictionary
c432	52	524	10	0.995422	0.910887
c499	54	758	10	0.997761	0.982002
c880	52	942	10	0.997370	0.960276
c1355	86	1574	11	0.998026	0.976322
c1908	121	1879	11	0.997644	0.954416
c2670	111	2747	12	0.998271	0.935137
c3540	154	3428	12	0.998390	0.879032
c5315	119	5350	13	0.999314	0.945444
c6288	32	7744	13	0.999749	0.998148
c7552	207	7550	13	0.999307	0.952915
Average				0.998125	0.949458

Table 4. Comparison of the Diagnostic Resolutions

of partitioning the test set and XORing the responses in each partition promises significant reduction in dictionary size, particularly for large industrial circuits, where such improvements are sorely needed.

6. Conclusion

A method is proposed to deal with the exceedingly large size of the fault dictionary. The test vector set, generated by an ATPG, is partitioned into blocks and a combined signature is created for the test vectors in each block. The scheme aims at providing high diagnostic resolution with a relatively small number of combined signatures.

Instead of storing pass-fail information for each test vector, a combined signature is stored for each test vector partition. Consequently, if high diagnostic resolution can be provided with a small number of partitions, a considerable decrease in the dictionary size can be achieved.

An algorithm is presented to partition the test vector set and the pass-fail responses of the test vectors in each partition are XORed to provide the combined response for the partitions. The proposed method is applied to the circuits in the ISCAS85 benchmark set. The experimental results indicate that orders of magnitude reductions in the dictionary size can be achieved at the cost of a typically insignificant decrease in diagnostic resolution. The enormous decrease in the dictionary size for the larger circuits in the experiments is highly promising in regards to the applicability of the proposed method on large industrial circuits.

References

[1] M. Abramovici, M. A. Breuer, and A. D. Friedman. *Digital Systems Testing and Testable Design*. IEEE Press, 1990.

[2] V. Boppana and W. K. Fuchs. Fault dictionary compaction by output sequence removal. In *Proceedings of the 1994 IEEE International Conference on Computer Aided Design*, pages 576–579, 1994.

[3] V. Boppana, I. Hartanto, and W. K. Fuchs. Full fault dictionary storage based on labeled tree encoding. In *Proceedings of 14th VLSI Test Symposium*, pages 174–179, 1996.

[4] F. Brglez and H. Fujiwara. A neutral netlist of 10 combinational benchmark circuits and a target translator in FORTRAN. *ISCAS*, June 1985.

[5] P. Camurati, A. Lioy, P. Prinetto, and M. S. Reorda. Diagnosis oriented test pattern generation. In *Proceedings of the European Design Automation Conference*, pages 470–474, 1990.

[6] B. Chess and T. Larrabee. Creating small fault dictionaries. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 18(3):346–356, March 1999.

[7] M. Garey and D. S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. Freeman, 1979.

[8] D. B. Lavo and T. Larrabee. Making cause-effect cost effective: Low-resolution fault dictionaries. In *2001 IEEE International Test Conference*, pages 278–286, 2001.

[9] H. K. Lee and D. S. Ha. *On the Generation of Test Patterns for Combinational Circuits*. Technical Report 12-93, Department of Electrical Eng., Virginia Polytechnic Institute and State University.

[10] H. K. Lee and D. S. Ha. HOPE: An efficient parallel fault simulator. In *29th ACM/IEEE Design Automation Conference*, pages 336–340, 1992.

[11] I. Pomeranz and S. M. Reddy. On the generation of small dictionaries for fault location. In *Proceedings of the 1992 IEEE International Conference on Computer-Aided Design*, pages 272–279, 1992.

[12] P. G. Ryan, W. K. Fuchs, and I. Pomeranz. Fault dictionary compression and equivalence class computation for sequential circuits. In *Proceedings of the 1993 IEEE International Conference on Computer-Aided Design*, pages 508–511, 1993.

[13] P. G. Ryan, S. Rawat, and W. K. Fuchs. Two-stage fault location. In *International Test Conference*, pages 963–968, 1991.